القياس النفسى

والاختبارات النفسية

الأستاذ الدكتور

عبد الحميد محمد على

أستاذ الصحة النفسية المساعد

رئيس قسم علم النفس التربوى

كلية التربية بالعريش - جامعة قناة السويس

الأستاذ الدكتور

فتحى عبد الحميد عبد القادر

أستاذ علم النفس التربوى

كلية التربية - جامعة الزقازيق

الدكتور

مراد على عيسى

مؤسسة طيبة

القياس النفسي والاختبارات النفسية

إعداد

الأستاذ الدكتور

فتحي عبد الحميد عبد القادر

أستاذ علم النفس التربوي

كلية التربية – جامعة الزقازيق

الأستاذ الدكتور

عبد الحميد محمد على

أستاذ الصحة النفسية المساعد

رئيس قسم علم النفس التربوي

كلية التربية بالعريش – جامعة قناة السويس

الدكتور

مراد على عيسى

الناشر

مؤسسة طيبة للنشر والتوزيع

7شارع علام حسين – ميدان الظاهر – القاهرة

ت – 0227876470 /0227867198

فاكس / 0227876471

محمول / 0112155522 – 0191848808- 0106242622

الطبعة الاولى 2010

فهرسة أثناء النشر من دار الكتب والوثائق القومية المصرية

على ، عبد الحميد محمد .

القياس النفسى والاختبارات النفسية / إعداد عبد الحميدمحمد على . – ط 1 . – القاهرة :

مؤسسة طيبة للنشر والتوزيع ، 2010

306 ص ؛ 24 سم .

تدمك : 9- 185- 431- 977- 978 .

1 – القياس النفسى

2 – الاختبارات النفسية

أ- العنوان 150,287

رقم الإيداع : 2010/2990

بسم الله الرحمن الرحيم

المقدمة

قام الإنسان بالقياس منذ أن وجد علي ظهر البسيطة ، فعندما كان يفصل ثوباً لنفسه كان يضع جلد الحيوان علي جسمه ، وما زاد عن طوله يقصه ، ويستخدم الشبر والميتر والخطوة والذراع والحبل والعصا في القياس ، كما أن الأطفال يقيسون المسافات خلال لعبهم بالقدم ، وبقي الحال علي هذا المنوال وقتاً طويلاً حتى ظهرت حالة الإنسان إلي المقايضة او المبادلة فنشأت مشاكل عديدة سببها اختلاف المقاييس من شخص لأخر ، ونشأت الحاجة إلي مقياس موحد متساوي الوحدات لا يتأثر برغبات الإنسان ، فظهر المتر واللتر والكيلو جرام وغيرها .

أن استناد القرارات إلي إجراءات صحيحة وصادقة أمر هام وحيوي ، هذا علي الرغم من أن القرارات لا تتساوي في أهميتها ، فبعضها أهم من البعض الأخر ، ولكن كلاً منهما يؤثر علي نحو مباشر أو غير مباشر في التلاميذ غير أن كل قراراً تربوي ، ينبغي أن يتم علي أساس عقلاني وموضوعي بأكبر درجة ممكنة وأن يستند إلي أفضل البيانات المتوافرة ، وتبلغ جودة البيانات التي جمعت خلال عملية التقويم مبلغ القياسات التي تعتمد عليها ، أن القياس هو مقدار ما يملكه شخص معين من سمه ، والقياس يوفر وصفاً أكثر موضوعية للسمات وييسر المقارنات ، وبدلاً من أن نقول هذا الطفل قصير بالنسبة لسنه يمكن أن نقول يبلغ طول قامة الطفل (130) سم ، وبدلاً من ان نقول أن (س) اكثر ذكاءً من (ص) نقول ان نسبة ذكاء (س) هي (120) ، ونسبة ذكاء (ص) هي (90) . والعبارة الرقمية اكثر دقة وموضوعية من العبارة اللفظية وأقل قابلية للتفسيرات منها .

الاختبارات النفسية هى أدوات الحصول على المعلومات والتى يمكن أن يمدنا بها الاختبار، وعلى المرء أن يحتفظ بهذه المعلومة الضرورية فى عقله ، وإلى أداة من الممكن أن تصبح أداة للمنفعة أو الإفساد ، وهذا يعتمد على كيفية استخدامها.

ولقد نمت الاختبارات على مدى خطوات متزايدة ومستمرة ، كما إنها تسهم بشكل فعال فى مساحات أكبر وأكبر فى الحياة اليومية. ولكن هذا النمو رافقه بعض الآمال الغير واقعية وسوء الاستخدام؛ فمستخدمو الاختبار فى حاجة لأن يعرفوا كيف يقوّموا الاختبار ، ما مدى جودة الاختبار بالنسبة للأغراض الخصوصية التى استخدم من أجلها؟ ما نوع المعلومات التى يمكن أن يمدنا بها الاختبار عن الشخص المفحوص؟ كيف يمكن أن تكمل النتائج الشبكة من المعلومات التى نحصل عليها للوصول إلى قرارات فاعله؟

فهذا الكتاب وضع للإجابة عن هذه الأسئلة ، وهو لم يصمم للمتخصصين فى الاختبارات فقط ، ولكن أيضا لعامة طلاب علم النفس. فاليوم نحتاج لبعض المعرفة الأساسية عن الاختبارات ، ليس فقط من قبل المصممين أو الفاحصين ولكن أيضا لأى شخص يستخدم نتائج الاختبارات كأحد مصادر المعلومات للوصول لقرارات عن الفرد نفسه وعن الآخرين.

و الله نسأل أن يجعل هذا العمل فى ميزان حسناتنا و ينفع به أمتنا

المؤلفون

الباب الأول

القياس النفسي

الفصل الأول

" المدخل للقياس النفسي "

مقدمة :

قام الإنسان بالقياس منذ أن وجد علي ظهر البسيطة ، فعندما كان يفصل ثوباً لنفسه كان يضع جلد الحيوان علي جسمه ، وما زاد عن طوله يقصه ، ويستخدم الشبر والمتر والخطوة والذراع والحبل والعصا في القياس ، كما أن الأطفال يقيسون المسافات خلال لعبهم بالقدم ، وبقي الحال علي هذا المنوال وقتاً طويلاً حتي ظهرت حالة الإنسان إلي المقايضة أو المبادلة فنشأت مشاكل عديدة سببها اختلاف المقاييس من شخص لأخر ، ونشأت الحاجة إلي مقياس موحد متساوي الوحدات لا يتأثر برغبات الإنسان ، فظهر المتر واللتر والكيلو جرام وغيرها .

أن استناد القرارات إلي إجراءات صحيحة وصادقة أمر هام وحيوي ، هذا علي الرغم من أن القرارات لا تتساوي في أهميتها ، فبعضها أهم من البعض الأخر ، ولكن كلاً منهما يؤثر علي نحو مباشر أو غير مباشر في التلاميذ غير أن كل قراراً تربوي ، ينبغي أن يتم علي أساس عقلاني وموضوعي بأكبر درجة ممكنة وأن يستند إلي أفضل البيانات المتوافرة ، وتبلغ جودة البيانات التي جمعت خلال عملية التقويم مبلغ القياسات التي تعتمد عليها ، أن القياس هو مقدار ما يملكه شخص معين من سمه ، والقياس يوفر وصفاً أكثر موضوعية للسمات وييسر المقارنات ، وبدلاً من أن نقول هذا الطفل قصير بالنسبة لسنه يمكن أن نقول يبلغ طول قامة الطفل (130) سم ، وبدلاً من ان نقول ان (س) أكثر ذكاءً من (ص) نقول ان نسبة ذكاء (س) هي (120) ، ونسبة ذكاء (ص) هي (90) .

والعبارة الرقمية اكثر دقة وموضوعية من العبارة اللفظية واقل قابلية للتفسيرات منها .

وقد ورد من عالم النفس الأمريكي " ثورنديك" قوله : إذا وجد شيئاً فأنه يوجد بمقدار ، فإذا كان بمقدار فمعني ذلك أننا نستطيع قياسه ، والقياس أضيق في معناه من التقويم ويعني غالباً جمع معلومات كمية عن موضوع القياس . ويلجأ المعلم إلي القياس في أحوال كثيرة ، فهو يقوم مثلاً لقياس مقدرة الطفل عند التحاقه بالمدرسة لأول مرة ومتي دخلها ، يقوم بقياس قدرته علي فترات للتأكد من استفادته من جهود المعلم وقياس مدي هذه الاستفادة ثم يعود فيقوم بقياس قدرته عند الانتقال إلي مرحلة دراسية أعلي أو إلي الحياة العملية ، وتتم عملية القياس في كل من هذه المراحل بواحدة من عدة وسائل منها : التقدير الشخصي ، والامتحانات والاختبارات والمقاييس العقلية والتحصيلية .

إن هدف القياس التربوي هو أن يبين مقدار ما يمتلكه شخص معين ، أو وحدة معينة من خاصية بعينها هل هو قليل أو متوسط ، أو كثير وان يعبر عن ذلك تعبيراً كمياً .

وكلمة قياس ليست مرادفة لتطبيق اختبار . ذلك أنه يمكن جمع البيانات والمعلومات من خلال عمليات كالملاحظة ومقاييس التقدير المتدرجة .

أن القياس الصادق في مجال التربية ليس عملاً سهلاً ، والمشكلة الأساسية هي أنه باستثناء قياس الخصائص الجسمية كالطول والوزن تعتبر جميع القياسات غير مباشرة ، ولا توجد مقاييس مترية لقياس سمات مثل الذكاء والتحصيل والاتجاهات ، ولذلك فإن القياس نوعان هما :

. قياس مباشر محدد : ويحدث عندما نقيس طول قامة إنسان أو ارتفاع بنائه أو قياس حجم مربع .

ـ قياس غير مباشر : ويحدث حين نقيس تحصيل التلاميذ أو استعداداتهم وقدراتهم العقلية بالاستجابة لمواقف معينة .

مفهوم القياس :

يمكن تعريف القياس إحصائيا بأنه تقدير الأشياء والمستويات تقديراً كمياً ، وفق إطار معين من المقاييس المدرجة ويتضمن القياس علاوة علي التقدير الكمي للظواهر عملية مقارنة نتيجة القياس بغيرها .

ويعرف " كامبل " (1925) القياس علي انه تمثيل للصفات أو الخصائص بأرقام ، كما يعرفه " ستيفنز " (1951) بأنه عملية تحديد أرقام لأشياء أو أحداث وفقا لقوانين أي لنقيس شيئاً علينا أن قوم بعمليات نقارن الشئ بمعيار أو بمقياس معين حسب قواعد معينة . ويري " جيلفورد" (1954) ان القياس وصف للبيانات أو المعطيات بالأرقام ، ويري " تنالي " (1970) أن القياس يتكون من قواعد استخدام الأعداد بحيث تدل علي الأشياء بطريقة تشير إلي كميات من الخاصية وقد ورد في قاموس " وبستر " أن القياس هو التحقق بالتجربة أو الاختبار من المدى أو الدرجة أو الكمية أو الأبعاد أو السعة بواسطة معيار ، ويتضح أن جميع التعريفات تتضمن التعبير عن النتيجة بأرقام أي أن القياس عملية كمية وأن نتائجه يعبر عنها بأرقام . اما " اوليرة " (1985) فيعرف القياس بأنه العملية التي تتحدد بواسطتها كمية ما يوجد في الشئ من الخاصية أو السمة التي نقيسها . وإذا اعتبرنا أن لغة الكم هي جوهر القياس بصفة عامة سواء بالنسبة للأشياء أو الأشخاص ،فأن القياس يمكن تعريفه بأنه " قواعد استخدام

الأعداد بحيث تشير إلي الأفراد أو الأشخاص ، وتدل علي كميات من صفة أو خاصية أو سمة ". ويتضمن هذا التعريف ما يأتي :

أ- أن إجراءات استخدام الأعداد في القياس يجب أن تصاغ صياغة صريحة وأن يعبر عنها بوضوح بحيث تصبح من قبيل القواعد ، صحيح أن بعض القواعد قد يكون في بعض الأحيان علي درجة كبيرة من الوضوح بحيث لا تتطلب صياغة تفصيلية كما هو الحال في استخدام المتر في قياس الأطوال إلا أن هذه الأمثلة هي الشذوذ في ميدان القياس وليست القاعدة في العلم ، فعند قياس مقادير العناصر المختلفة التي تتكون منها المركبات الكيميائية يتطلب الأمر إجراءات متعددة ليست واضحة بذاتها كما هو الحال في المتر ، من المؤكد أن قواعد قياس كثير من الخصائص أو السمات النفسية والتربوية ليست واضحة بذاتها كما هو الحال في قياس الذكاء مثلاً .

ب- أن الأعداد التي تستخدم في القياس لا تقيس الأشياء أو الأشخاص وإنما تشير إليها ، وما يخضع للقياس هو الصفات أو الخصائص أو السمات وهذه الخصائص أو السمات هي تجريدات بها نعرف الأشياء و الأشخاص .

ج- أن الأعداد في إشارتها للأشياء والأشخاص تدل علي كميات من خصائص أو سمات هذه الأشياء وهؤلاء الأشخاص ، وفكرة الكم تعني مقدار ما يوجد في الشئ أو الشخص من الخاصية أو السمة .

أخطاء القياس:

مهما حولنا أن نكون دقيقين في القياس فلابد أن نقع في أخطاء بعضها خارج عن أرادتنا ومن الضروري أن نعي هذه الأخطاء قبل قيامنا بالقياس كي نحد منها وكذلك لنجرد منها درجة التلميذ بواسطة الطرق الإحصائية ثم لنأخذها بعين الاعتبار عند تفسير الدرجة وهذه الأخطاء ناجمة عن :

1. عدم حساسية أدوات القياس :

بعض أدوات القياس أكثر دقة من بعضها الأخر، فالميزان الحساس أكثر دقة من الميزان الزنبركي ، فالأول يعطينا الوزن بأجزاء الجرام والثاني يعطينا الوزن بالكجم فقط ولا يقيس الجرامات الزائدة ، وفي ميدان التربية يعتبر الامتحان الموضوعي أكثر دقة من الامتحان المقالي ، كما أن هناك فرقاً بين امتحان موضوعي يصممه فاحص مدرب وآخر يضعه فاحص لا يدري عن القياس شيئاً .

2. عدم ثبات الظاهرة المقاسة :

معظم ما يقاس في ميدان التربية وعلم النفس ديناميكي أي متقلب متغير فالشخص الذي نقيس أداءه هو كل ساعة في شأن ، فهو الآن سليم معافى ولكنه قد يشعر بالتعب أو الملل بعد قليل أو يفقد اهتمامه وتنخفض فاعليته ويتأثر أيضاً بدرجة الحرارة والتهوية والإضاءة وبقية المتغيرات البيئية ، ولذا فأننا نحاول أن نقترب من الحقيقة بأخذ عدة قياسات ثم استخراج متوسطها واعتباره أقرب ما يكون لأدائه .

3. خطأ الملاحظة :

تختلف دقة الملاحظة من شخص لأخر فلو طلبنا من عدة معلمين قياس تحصيل تلميذ في مادة معينة فسيعطوننا درجات مختلفة مع أن أداة القياس لم تتغير وكذلك التلميذ موضوع القياس ، وهذه الفروق في القياسات ممكنة الحدوث لدي الشخص نفسه إذا قام بالقياس عدة مرات .

طبيعة القياس النفسي والعوامل المؤثرة فيه :

يختلف القياس باختلاف الذاكرة المقاسة ففي الفلك والفيزياء والكيمياء يكون القياس أدق ما يمكن ولا تقبل بوجود اخطاء مهما صغرت أما إذا انتقلنا إلي ميدان النبات والأحياء فان قياسنا يصبح اقل دقة لأن هناك عوامل مؤثرة اكثر عدداً وقابليتها للتغير أشد وأعظم ، وتكون هذه العوالم الاجتماعية والسلوكية ولذا نكون اكثر عرضه للخطأ ويمكن أن نتجاوز عنه إذا كان لا يزيد عن (5٪) إلا انه يجب ان ننتبه إلي هذه الأخطاء كي نستعمل أدوات القياس ونفسر نتائجه تفسيراً صحيحاً ، وهذا يقتضي أن نلم بخصائص القياس النفسي التالية :

1) القياس النفسي كمي .

2) القياس النفسي قياس غير مباشر ، فنحن مثلاً لا نقيس الذكاء بعينه وإنما نستدل عليه من أداء التلميذ .

3) في كل قياس نفسي يوجد خطأ ما ، وهذا يعني أنه علينا أن نكتشف هذا الخطأ بالطرق الإحصائية ثم نزيله قبل استخدام النتائج أو تفسيرها.

4) القياس النفسي نسبي أي غير مطلق فلا وحدات تربوية أو نفسية لدينا مماثلة للسنتيمتر أو الصف وعليه لا يمكن تفسير نتائج القياس إلا بمقارنتها بمعيار أو مستويات مشتقة من أداء مفحوصين أو أداء عينات مماثلة لعينة المفحوصين .

5) علي الرغم من أن وحدات القياس النفسي توصف بأنها غير متساوية فعلاً فأن تقدماً قد حدث في أيجاد وحدات تتصف بالثبات والتساوي مثل الدرجات المعيارية .

6) الصفر في المقاييس النفسية صفر عرفي فهو ليس بصفر حقيقي أي لا يدل علي عدم وجود الشئ ، فإذا حصل طالب علي الدرجة صفر في امتحان تاريخ

فلا يعني ذلك أنه لا يعرف شيئاً في التاريخ وإنما لا يعرف شيئاً بالنسبة لهذه العينة من الأسئلة لو استبدلناها بأسهل منها أو حتي بغيرها لوجدنا انه يعرف الإجابة عن بعضها وبناء عليه يجب أن نعتبر الصفر التربوي أو النفسي يدل علي درجة ما ولا يدل علي عدم وجود الشئ ، ولما كان هذا الصفر يمثل نقطة أصل أو نقطة مرجعية فيستحيل أن نتخذ متوسط الدرجة كنقطة استناد بدلاً منه ونقيس درجة الطالب ببعدها عنه عندما نريد تفسيرها .

أما العوامل المؤثرة في القياس النفسي فهي :

الشئ المراد قياسه أو السمة المراد قياسها : هناك أشياء تقاس بطريقة مباشرة كما يحدث حين نقيس طول سلك أو قياس أطوال التلاميذ عند الكشف الطبي عليهم ، ولكن السمات التي نقيسها في دراسة الاستعدادات العقلية والسمات المزاجية تقاس بطريقة غير مباشرة ، ولذلك تختلف المقاييس ليس فقط في طبيعتها بل وفي دقتها أيضاً فليس هناك شك في أن القياس المباشر أسهل وأدق من القياس الغير مباشر .

طبيعة الهدف من القياس :

تؤثر أهداف القياس في الطريقة التي نستخدمها في القياس ، فعندما يكون الهدف من القياس عمل تقويم عاجل لسمة معينة أو لتحصيل التلاميذ في خبرة معينة اختيرت الطرق التي تناسب هذا الهدف وحين يكون الهدف عمل تقويم شامل ودقيق لظاهرة معينة رسمت خطة دقيقة للقياس .

طبيعة العملية أو السمة المقاسة :

يتأثر القياس بالسمة المقاسة فبعض السمات يمكن التحكم فيها وقياسها بدقة لا بأس بها كما هو الحال في الذكاء والاستعدادات العقلية ، بينما يصعب

التحكم في بعض السمات المزاجية ، كما يصعب تحديدا تحديداً دقيقاً وتصميم المواقف التي تمثلها تمثيلاً صحيحاً بسبب تعقدها وتأثرها بعوامل ذاتية متعددة من الصعب استبعادها .

. مدي خبرة من يقوم بالقياس بأصول هذا العلم ومناهجه وأساليبه ومدي تدريب الذي يقوم بالقياس وجمع الملاحظات .

. عوامل أخري متعلقة بطبيعة الظاهرة المقاسة من جهة وبطبيعة المقاييس من جهة أخري .

هدف القياس وأهميته في التربية والتعليم :

أن الهدف من القياس في التربية والتعليم هو الكشف عن الاختلافات الفردية بأنواعها المختلفة ، فهناك اختلاف في الذكاء وفي سرعة التعلم والاستجابة والمهارة وغيرها ، إذاً أنه لولا وجود هذه الاختلافات الفردية ما كانت هناك حاجة إلي القياس . وتتلخص انواع **الاختلافات الرئيسية في :**

أ- **الاختلاف بين الأفراد :** والغرض من قياس هذا النوع هو مقارنة الفرد بغيره من افرا صفة أو عمره أو بيئته في ناحية من النواحي النفسية أو التربوية أو المهنية .

ب- **الاختلافات في الفرد نفسه :** ويهدف قياس هذا النوع من الاختلافات إلي مقارنة النواحي المختلفة في الفرد نفسه لمعرفة نواحي القوة والضعف في الفرد بالنسبة إلي نفسه ، وتوجيهه تربوياً ومهنياً .

ت- **الاختلاف بين الجماعات :** فالجماعات تختلف في خصائصها ، وقياس مثل هذه الاختلافات يفيدنا في دراسة سيكولوجية الجماعات وخصائص النمو ، ودراسة العوامل التي قد تكون مسئولة عن هذه الفروق .

ث- **الاختلاف بين المهن المختلفة :** تتطلب المهن المختلفة مستويات مختلفة من القدرات والاستعداد والسمات ، وقياس هذه الاختلافات يفيدنا في انتقاء المهنة وفي التوجيه المهني .

وتتحدد أهمية القياس في التربية والتعليم فيما يلي :

1. **دراسة التخلف الدراسي وأسبابه :**

عن طريق الاختبارات النفسية التي تكشف عن التخلف الدراسي وأسبابه . وهذه الاختبارات تعطي صورة كاملة عن جوانب حياة التلميذ النفسية .

2. **التوجيه والإرشاد التربوي :**

تساعد الاختبارات النفسية والتربوية التلميذ فى اختيار دراسة تناسبه بعد معرفته بقدراته واستعداداته وميوله وسماته .

3. **أختيار فروع التعليم المناسبة :**

تزود الاختبارات المدارس والتلميذ بصورة دقيقة عما تتطلبه دراسة معينة من تفوق فى مواد أخرى ، كما تساعد الاختبارات النفسية التلميذ على فهم مستواه فى مواد وقدرات معينة تؤهله لدراسات تالية أو تجعله يكف عن متابعة دراسة معينة لأنها تتطلب ما لا يتوفر لدية من قدرات .

4. **تخطيط وتنفيذ البرامج الشاملة للاختبارات :-**

أن الاختبارات المتبعة فى التربية والتعليم تحتاج إلي تقييم ومتابعة لمعرفة مدي صلاحيتها ، ولكي تقوم أجهزة التربي والتعليم بوظائفها لكي نحسن تقدير التلاميذ ونقيم المدرسة والإدارة .

أنواع المقاييس :

تختلف المقاييس النفسية والتربوية عن بعها من حيث نظام القياس أو اسلوبه ودقته وفيما يلي هذه المقاييس مرتبة من ابسطها إلي أعقدها .

المقياس الاسمي Nominal Scale :

ان أبسط خطوات القياس هي تصنيف الأشياء أو الأشخاص إلي فئات فالطلاب يصنفون إلي راسب ، ناجح أو إلي ابتدائي ، أعدادي ، ثانوي ، جامعي . أو إلي ذكور إناث . هذا ويمكن الاستعاضة عن السماء بأرقام فالذكورة مثلاً تعطي الرقم واحد وصفة الأنوثة تعطي الرقم (2) او رقم أخر بحيث ينطوي تحت لوائه كل من يتصف بصفة الأنوثة من مجتمع الدراسة أو بحيث يدل الرقم علي أن جميع الأفراد متساوون في خاصية لم يمتلكها غيرهم .

وتستعمل الأعداد في هذا المقياس وكذلك الرموز للدلالة علي الأشخاص . مثلاً ، لكل عضو في فريق كرة القدم رقم يدل عليه ، كما أن لكل تلميذ رقم جلوس يشير أليه بحيث يحل العداد محل الأسماء .

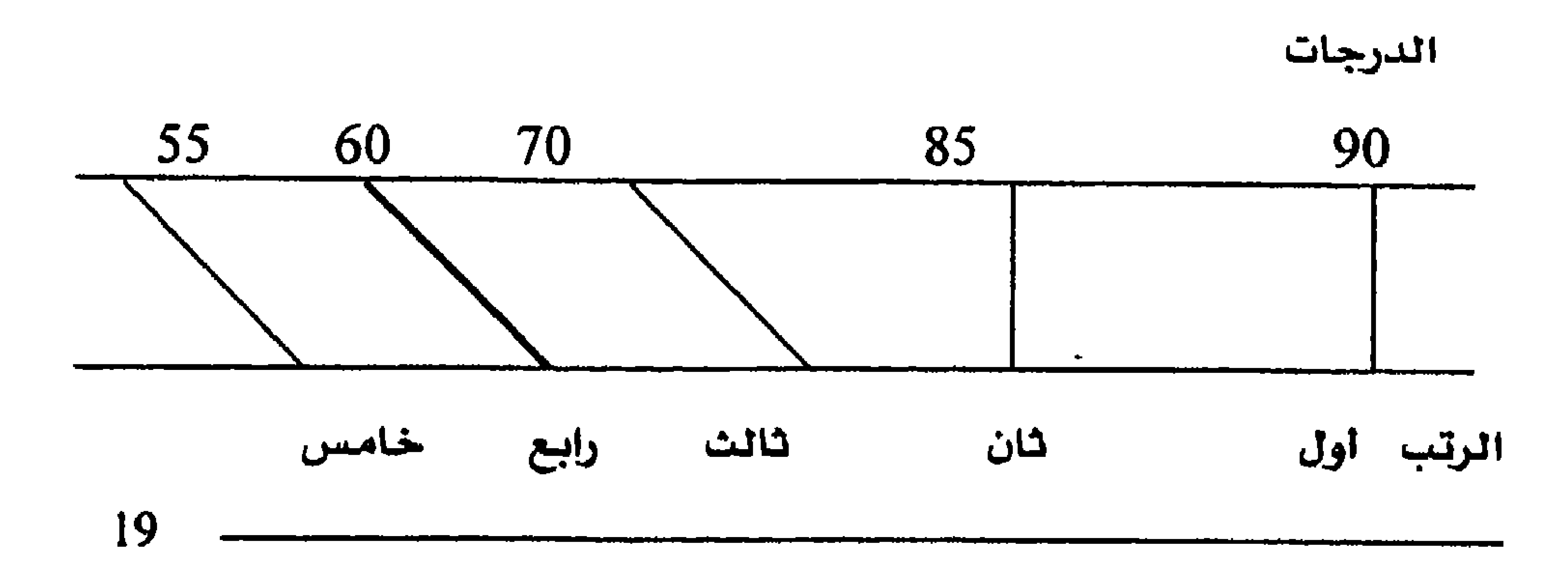
ولا تدل الأعداد في هذا المقياس ، إذا استعملت كما سبق علي كميات إذ ليس لها مضمون كمي يساوي مقدار ما يوجد بالشئ من صفة أو خاصية ، وإنما تدل علي معني كيفي أي علي نوع المعدود ، العملية الحسابية التي يمكن أن نستفيدها من هذا المقياس هو العد بعد التصنيف وما يترتب عليها فإذا كان أكبر رقم يحمله لاعب كرة القدم (11) فمعنى ذلك ان هناك احد عشر لاعباً ، وعندما نصنف الطلبة إلي ذكور وإناث ، إذاً أعطينا الذكور الرقم (1) ، وصفة الأنوثة الرقم (2) فأن هذين الرقمين يدلان علي جنس الفئة كما انه باستطاعتنا عد افراد الفئة . اي أن المقاييس الاسمية تستخدم لغة العدد دون أن تدل علي كميات .

مقياس الرتب Ordinal Scale :

عندما نرتب درجات التلاميذ ترتيباً تنازلياً أو تصاعدياً ثم نستخرج رتبها فأننا نستعمل مقياس الرتب . كذلك حين نرتب المتسابقين حسب وصولهم إلى خط النهاية فأن المتسابق (أ) يمتلك من القدرة علي العدو أكثر من البقية إذا وصل إلى خط النهاية . أما من وصل بعده فيمتلك من هذه القدرة أكثر من بقية المتسابقين ولكنه يمتلك منها مقداراً أقل مما يمتلكه لمتسابق (أ) أما بالنسبة للأخير فهو يمتلك منها مقداراً أقل مما يمتلكه اي فرد أخر . هذا المقياس يدل علي أن الشخص يمتلك من السمة المقاسة أكثر أو أقل مما يمتلكه أخر أو آخرون ولكنه لا يدل علي مدي أو مقدار ما يمتلكه كلاً منهم .

وهذا المقياس شائع الاستخدام في ميادين علم النفس والتربية ولاسيما حيث يتعذر القياس الموضوعي كما في التربية البدنية والموسيقي والخط والغناء أو عند دراسة المواقف والميول ولذا فأننا نعطي رتباً لأتن تحديد الكم أو المقدار أي مدي الجودة والأفضلية بناء علي أسس موضوعية دقيقة أمر متعذر .

ونظراً لأن وحدات مقياس الرتب متساوية ظاهرياً وغير متساوية فعلياً ، إذاً أن الفروق المتساوية في الرتب لا تدل علي فروق متساوية في الدرجات الخام كما في الشكل التالي . فأننا لا نستطيع جمعها أو طرحها أو قسمتها أو استخراج متوسطها وانحرافها المعياري . ولكننا نستطيع استخراج الوسيط إلا انه يتوقف علي التوزيع التكراري وليس علي الدرجات ، كما يمكن أن نحو الرتب إلي رتب مئينية أو نستخرج لها معامل ارتباط الرتب.



مقياس المسافة Interval Scale :

هـذا النـوع مـن المقـاييس أدق المقـاييس إذا انـه يتمتـع بوحـدات متـساوية لتمكننـا مـن أن نحـدد مـا إذا كـان شئ يـساوي شيئاً آخـر أو اكـبر أو اصـغر منـه ومدي الكبر أو الصغر اي بكم يزيد او يقل عنه في السمة المقاسة وبعبارة أخري وعلي العكس من مقياس الرتب ، فأن المسافات المتساوية علي هذا المقياس تدل علي مقـادير متـساوية مـن الخاصية التي نقيسها ولـذا نستطيع جمع هـذه المسافات او طرحها ولكن ينقص هذا المقياس وجود صفر حقيقي او مطلق كالصفر الرياضي الـذي يـدل علـي عـدم وجـود شئ ممـا يهـدف إلـي قياسـه . ومـن الأمثلـة علـي هـذا المقياس كل من المقاسين الفهرنهيتي والمئوي وهنا يجب أن ننبه إلي أن الصفر في المقيـاس المئـوي مـثلاً صـفر نـسبي أو اعتبـاطي أو عـرفي اي أنـه لا يـدل علـي غيـاب السمة المقاسة أي علي عدم وجـود الحـرارة وإنمـا يـدل علـي النقطـة أو الدرجـة الـتي يتجمد عندها الماء وهذا الصفر مماثل للصفر المستعمل حالياً من قبـل المعلمين فعلي الرغم من أنهم يريدون أن يبينوا أن الطالب الحاصل علي الصفر في مادة ما لا يمتلك اية معلومـات إلا أنـه منطقيـاً يعـرف شيئاً مـن هـذه المادة وهـذا الشئ لم يقـسه المقيـاس أو الامتحـان والـصفر فـي مقيـاس المسافة أو المقيـاس المئـوي مـثلاً ينظر إليه كنقطة مرجعية تحدد بواسطتها ما إذا كانت درجة الحرارة فوق درجة تجمد الماء أم أقل منها .

وفي القيـاس التربـوي جـرت العـادة علـي أن تتخـذ الوسـط الحـسابي لتوزيـع الدرجات نقطة أصل أو استناد ولذا ومن ثم نرمـز إليـه بـصفر أو تطـرح منـه كميـة ثابتة متساوية ثم نحدد بعد العلامـات عنـه بمقـدار زيادتهـا او نقـصانها عنـه اي ان درجة الطلب تصبح مساوية لانحرافها عن المتوسط . فإذا كـان متوسـط الامتحـان (60) وكانـت درجـات بعـض الطلبـة تـساوي (75 ، 65 ، 60 ، 48) فإنهـا تـصبح (15 ، 5 ، صفر ، - 12). اي اننـا قـسنا درجـة الطالـب بمتوسـط الـدرجات وعـادة

نقيس بعد الدرجة عن المتوسط بوحدة الانحراف المعياري فإذا كان هذا يساوي خمس درجات وقسمنا عليه الانحرافات السابقة فإن العلامات تصبح (3 ، 1 ، صفر ، - 2.4) والإشارة السالبة تعني أن العلامة تحت الوسط بمقدار يساوي (- 2.4) وحدة معيارية .

وعلي العكس من المقاييس السابقة يمكن أن نستخدم مع هذا المقياس عمليات الجمع والطرح والضرب دون أن نغير جوهر العلاقة بين العلامات أما القسيمة فلا نستطيع القيام بها لأنها تفترض وجود الصفر المطلق ولتوضيح ذلك نفترض أن درجة الحرارة في اسوان (30 درجة)، وأنها في الإسكندرية (15 درجة) فهل درجة الحرارة في أسوان ضعفها في الإسكندرية ؟ الجواب لا إذاً يجب ان نأخذ بعين الاعتبار مدي بُعد كلاً منهما عن الصفر المطلق وهو (273 + 30 =303)في الحالة الأولي ، (273 + 15 = 288) في الحالة الثانية . ومنه نري أن درجة الحرارة في أسوان وهي (303) ليست ضعف (288) وهي درجة حرارة الإسكندرية .

وبناء علي ما سبق نحن لا نستطيع أن نقول أن الذكاء أو تحصيل فلان ضعف ذكاء أو تحصيل علان نظراً لعدم وجود صفر حقيقي يمثل انعدام الذكاء او التحصيل بل أن الذكاء والتحصيل لا يمكن أن ينعدما لدي أي فرد من الأفراد فهو يتمتع بقسط منه الذكاء أو التحصيل مهما كان ضئيلاً وهذا يعني أنه من المستحيل أن يكون هناك صفر حقيقي في ميدان علم النفس أو التربية .

مقياس النسبة Ratio Scale :

يتميز هذا النوع من المقاييس بأن له وحدات متساوية وصفراً مطلقاً ، ومثال عليه مقياس " كالفن Kelvin " وهذا صفره المطلق نقطة تنعدم عندها الحرارة وتساوي (- 273 درجة) كما ان جميع مقاييس الطول والوزن والسعة أو

الحجم هي من هذا القبيل فالمتر له صفر ووحدات متساوية وبواسطتها نستطيع أن نقول أن طول نقطة ما صفر وأن طول غرفة يساوي (6) امتار ، وطول آخر (12) متر وأن طول الثانية يساوي ضعف طول الأولي .

هذا النوع من المقاييس مستعمل علي نطاق واسع في ميدان العلوم ، ويندر أن يستعمل في نطاق التربية وعلم النفس اللهم إلا عندما نقوم بالقياس المباشر لأطوال التلاميذ وأوزانهم وكذلك عندما نقيس أزمنة الرجع والتعلم .

جدول يبين أنواع المقاييس وخصائصها

المقياس	الخاصية	العملية التجريبية الأساسية المؤدية إليه	المعالجة الإحصائية
الاسمي أو التصنيفي	المساواة	التصنيف إلي فئات بحيث يتساوى أفراد الفئة في صفه ما : ذكر، أنثي ، ناجح ، راسب .	عدد الحالات في الفئة . المنوال . الوسيط
الترتيبي	التساوي . الترتيب	تحديد أكبر من ، أصغر من . أكثر أقل . إدخال فكرة الزيادة دون تحديد كميتها	الوسيط . المئينيات . معامل ارتباط الرتب
المسافة	التساوي . الترتيب . وحدات متساوية	تدريج المقياس بحيث تتساوي وحداته لتبين الزيادة وكميتها مثل مقياس الحرارة	جميع العمليات الحسابية عدا القسمة . المتوسط . الانحراف المعياري . معامل ارتباط بيرسون . اختبارات الدلالة "ت" ، "ف"
النسبي	التساوي . الترتيب . وحدات متساوية . صفر حقيقي	تحديد صفر حقيقي مثل طول نقطة هندسية ؛ تحديد تساوي النسب	جميع العمليات الحسابية مفتوحة والمعالجات الإحصائية المختلفة ممكنة

الفصل الثاني

" أدوات القياس "

مقدمة :

علي الرغم من تعدد وتنوع أدوات القياس التي يمكن للمعلم أن يستخدمها فيجمع البيانات عن التلاميذ. إلا أن أسلوب وطريقة الاختبار هي أكثر هذه الأدوات أو الأساليب شيوعاً واستخداما بين المعلمين في المدارس ، وخاصة الاختبارات المدرسية التي تستخدم علي مدار العام الدراسي لغايات مختلفة كالتشخيص أو التصنيف أو التوجيه أو التقويم ، ومع ذلك تبقي أدوات وأساليب القياس الأخرى كالاستخبارات واستمارات الملاحظة والمقابلات الشخصية ومقاييس التقدير والأساليب الاسقاطية ، أدوات وأساليب هامة في القياس تزودنا بمعلومات وبيانات مهمة عن التلاميذ في جوانب ومواقف معينة .

وسنعرض بعض أساليب وأدوات القياس التي تناسب المجال التربوي ، مع التركيز علي الاختبار كأداة رئيسية شائعة الاستخدام في مدارسنا .

أولاً : الاختبارات :

تعريف الاختبار :

يقصد بالاختبار أنه أداة أو وسيلة أعدت بطريقة منظمة من مجموعة من المثيرات لتقدير خاصية سلوكية محددة لدي التلميذ بالتعبير عنها في صورة كمية أو رقمية ويتضمن هذا التعريف أركاناً أساسية أربعة هي : التقدير الكمي أو الرقمي كنتائج الاختبار ، والخاصية أو السمة المراد قياسها ، والمثيرات أو المفردات التي يتكون منها الاختبار والطريقة المنظمة أو الإجراءات المتبعة في إعداد الاختبار ونتناول فيما يلي هذه الأركان الأربعة :

التقدير الكمي أو الرقمي :

يجب أن يعطي الاختبار بيانات رقمية ، أو تقديرات كمية أو درجات عن الخاصية أو الشئ المراد قياسه فعند قياس تحصيل التلميذ في مادة ما نستخدم اختبار تحصيلي في هذه المادة كأداة لتحديد مقدار ما يتعلمه التلميذ منها ، وبالتالي فأن هذه الأداة تعطينا درجات لهذا التلميذ أو التلاميذ الآخرين علي نحو ييسر لنا اجراء مقابلة أو مقارنة كمية بين مستوي تحصيل هذا التلميذ ومستوي تحصيل التلاميذ الآخرين الذين تقدموا لنفس الاختبار .

الخاصية أو السمة المراد قياسها :

فيما يتعلق بالخاصية المراد قياسها بواسطة الاختيار ، فأن الاختبار لا يمثل كل جوانب السلوك المراد قياسه ، وإنما يكون الاختبار ممثلاً لعينة منه . فاختبار الحساب الذي يتكون من عشرين مسألة حسابية لا يمثل كل ما يعرفه التلميذ في هذه المادة من عمليات حسابية أو مسائل أو معادلات وإنما هو عينة منها . وان فقرات أو أسئلة الاختبار عينة منتخبة من السلوك المراد قياسه ، وأنها تعطي للتلاميذ كنوع من المثيرات لمعرفة مدي امتلاكهم منها .

المفردات التي يتكون منها الاختبار :

يجب أن تكون عينة ممثلة لكل السلوك المراد قياسه ، وأم تؤخذ هذه المفردات من هذا السلوك مباشرة ، وتتوقف قيمة الاختبار علي مدي الارتباط الحقيق بين أداء المفحوص له وبين أدائه في المواقف الأخرى المماثلة من حياته الواقعية .

الطريقة المنظمة أو الإجراءات المتبعة في إعداد الاختبار :

فهي تشير إلي القواعد والنظم المتبعة ف تكوين وتصميم الاختبارات ، فالاختبار التحصيلي وسيلة أو أجراء منظم لتحديد مقدار ما اكتسبه أو تعلمه

المتعلم ، والإجراء المنظم ملتزم به المعلمون والأخصائيين التربويين عند بناء الاختبارات ، تقديمه للتلاميذ للإجابة عليها وهي مرحلة التطبيق أو تصحيح إجاباتهم علي مثيراته أو أسئلته أو تفسير نتائجه وهي الدرجات التي حصل عليها هؤلاء التلاميذ ويعتبر الاختبار أداة رئيسية فعالة في يد المعلم ، ذلك أنه ييسر الحصول علي معلومات عن التلاميذ سواء عند التخطيط للتعليم الصفي ، أو عند تقييم الطرق أو الاستراتيجيات التي يستخدمها في التدريس ، وبالتالي فالاختبار يزود المعلم ببيانات عن أداء التلاميذ أو تحصيلهم تساعده علي معرفة مدي نموهم وتقدمهم أو تأخرهم ، ومن ثم تكوين فكرة عن مواطن القوة والضعف عند التلاميذ ومراعاة الفرق الفردية بينهم ، وزودهم بمعلومات مهمة عن ملائمة أو عدم ملائمة الأساليب والإجراءات التدريسية أثناء البرنامج التعليمي . وهذا يعني أن الاختبارات هي معينات تعليمية في يد المعلم إذا أحسن استخدامها فقد تكون الاختبارات المدرسية شفوية وخاصة في الصفوف الأولى من المدرسة الابتدائية ، حيث تكون الإجابة علي أسئلتها المطروحة من قبل المعلم بصورة شفهية ، وقد تكون الاختبارات تحريرية أو كتابية وفيها يستخدم التلميذ الورقة والقلم للإجابة علي الأسئلة ، وقد تكون الاختبارات عملية وفيها يقوم التلميذ بأجراء بعض الأعمال كالتجارب في العلوم أو أعمال فنية كالرسم والنحت والتمثيل والموسيقي وغيرها .

الاختبارات التحريرية :

تعتبر الاختبارات التحريرية أو الكتابية من أوسع أنواع الاختبارات انتشارا واكثرها شيوعاً بين المعلمين والأخصائيين التربويين ، وتعرف هذه الاختبارات أيضاً باختبارات الورقة والقلم ، وذلك لأن التلاميذ يجيبون علي هذه الاختبارات باستخدام القلم في الإجابة ، وتقسم هذه الاختبارات إلي اختبارات صفية واختبارات مقننه . فالاختبارات الصفية يضعها المعلم لتحديد مستوي تحصيل

التلاميذ ، وتقدم للتلاميذ في صورة اختبارات مقاليه أو اختبارات موضوعية أو خليط من الأسئلة المقالية والموضوعية . وأما الاختبارات المقننة تخضع لقواعد معينة عند تصميمها ، فضلاً عن ضرورة توافر خصائص محددة فيها كالصدق والثبات والموضوعية والتمييز والقابلية للاستعمال ، وهي اختبارات تستخدم في مجالات التحصيل الدراسي والاستعدادات العقلية ، بالإضافة إلي الشخصية والتكيف النفسي والاجتماعي . وسوف نعرض فيما يلي لاختبارات المقال والاختبارات الموضوعية .

أولاً اختبارات المقال :

وفيه يقوم التلميذ بكتابة موضع في الإجابة عن كل سؤال وهو النوع الغالب علي امتحانات مدارسنا والأكثر شيوعاً فيها .

واختبارات المقال ذات قيمة كبيرة في قياس القدرات الآتية :

1) القدرة علي اختبار المعلومات والمعارف التي تتصل بموضوع ما توضحه .
2) القدرة علي تنظيم المعلومات المختارة ، الأفكار المراد إظهارها في نسق متسلسل ووحدة منسقة .
3) القدرة علي التركيز في موضوع ما وعدم الخروج عنه واحكام دائرة التفكير .
4) القدرة علي استخدام جيد للأساليب اللغوية من حيث الكلمات المنتقاة .

وعلي الرغم من ذلك إلا أنها لها عيوب منها :

أ- عدم اتفاق المصححين علي الحقائق المطلوبة في الاختبار .

ب- ذاتية التصحيح .

ج- قلة عدد الأسئلة .

د- صعوبة التصحيح .

وعلي الرغم من تلك العيوب فأنها تعد وسيلة قياس جيدة مع توافر ما يلي :

1) أن يقوم بتصحيح جميع أوراق الإجابة المعلم الذي قام فعلاً بالتدريس

2) إذا أختلف المصححين يجب أن يكون المعلم الفعلي هو المرجع في معرفة المستويات ، الأهداف التي قصد تعليمها للتلاميذ .

3) يحاول المصحح أن يكون في حالة نفسية وظروف اجتماعية جيدة عند تصحيحه لجميع أوراق الإجابة .

4) عدم الاقتصار علي نوع (اختبارات المقال) بل يجب أن يشمل علي اختبارات موضوعية .

ثانياً الاختبارات الموضوعية :

يتكون الاختبار الموضوعي من عدد كبير من الأسئلة التي تأخذ أشكالاً مختلفة ، فقد تكون الأسئلة من نوع الصواب والخطأ ، أو الاختيار من متعدد أو من نوع المزاوجة أو من نوع التكملة أو أعادة الترتيب و في هذا النوع من الاختبارات نجد أن الدرجة التي يحصل عليها التلميذ في ورقة اجابته لا تختلف باختلاف المصحح .

مميزات الاختبارات الموضوعية :

أ- تغطي معظم أبواب المقرر .

ب- لا تتأثر بذاتية المصحح .

ج- سهلة التصحيح .

د- مناسبة للتلاميذ ذو الخط الردئ .

مواطن ضعف الاختبارات الموضوعية :

ليس نمن السهل وضع أسئلة الاختبارات الموضوعية فهي تحتاج إلي جهد كبير من المعلم ووقتاً طويلاً ، أي أنها صعبة الإعداد .

. لا تشجع علي التفكير ، كما أنها تحرم التلاميذ من إظهار القدرة علي التعبير .

. إنها تشجع علي التخمين في الإجابة ، وهذا غير مجد وإن أصاب أحياناً .

. أن هذه الاختبارات تقيس معرفة التلميذ للحقائق وتهمل اعتبارات أخري مثل معرفة درجة استيعاب المادة وحسن تطبيقاتها .

. أن هذه الاختبارات لا تتيح الفرصة للتلاميذ لإظهار قدراتهم .

أن الإجابة علي هذه الأسئلة تتطلب نوعاً آخر من التعلم مخالفاً تماماً للطريقة التي يطلب التلميذ بأتباعها في استخدام معلوماته في الحياة .

من أجل ذلك ينصح بعض المربين بالتوفيق بين القديم والحديث ، ويجب أن تكون الأسئلة في الاختبارات من النوع الواضح الخالي من اللبس وأن تستوفي الشروط الآتية :

- أن تشبع حاجات التلاميذ وقدراتهم المختلفة . ومن المستحب أن ترتب الأسئلة ترتيباً تصاعدياً من حيث الصعوبة او انها تساعد علي خلق الثقة في نفوس التلاميذ .

. بعد الانتهاء من التصحيح يجب ان تتأمل النتيجة والدرجات التي حصل عليها التلاميذ وبهذا يمكن الحكم علي ورقة الأسئلة أنها معقولة أو غير معقولة فالاختبار الجيد هو الذي تكون نتيجته متمشية مع منحني التوزيع الاعتدالي وإذا كان هناك عيباً في الامتحان أخذ المنحي شكلاً ملتوياً .

مراحل أعداد الاختبار التحصيلي الموضوعي :

هناك عدة مراحل يمر بها إعداد الاختبار التحصيلي من تجريب ، مراجعة ، تعديل قبل الصورة النهائية . وهي علي النحو التالي :

1) ابتكار فكرة الاختبار:

تعيد علي طبيعة المادة وظروف تدريسها والشروط المختلفة والعوامل التي تؤثر فيها ومستوي الدارسين ، إمكانيات المعلم والمؤسسة التربوية .

2) وضع مواصفات الاختبار :

يحسن انم يقوم واضع الاختبار بتوضيح فكرته توضيحاً دقيقاً وتمثل في :

. تحديد الأشياء المراد قياسها .

. شرح أنواع الوحدات التي سيتضمنها الاختبار .

. تحديد كل نوع من أنواع وحدات الاختبار .

. مدي الاختلاف في محتويات الاختبار وموضوع كلاً منهما .

. اقتراح الزمن المقرر .

. اقتراح الأسلوب الذي سيتم به تصحيح الاختبار .

3) المرحلة الاستطلاعية :

في هذه المرحلة يقدم في هذه المرحلة يقدم المخطط أولياً لهذا الاختبار ، ثم يقوم بتطبيقها علي أية مجموعة من الأفراد الذين ييسر استخدامه للتجربة . والغرض الرئيسي من تطبيق الاختبار هنا .هو مدي صلاحية وحدات الاختبار ولمعرفة الزمن الذي يستغرق في الاختبار،و الإحاطة بأوجه النقص .

4) الصورة الأولية :

وهي التي تأتي بعد التعديل حيث تصل إلي صورة صالحة للتطبيق علي مجموعة من نفس المجتمع الذي يوضع من أجلها الاختبار .... الغرض منها الحصول علي البيانات الإحصائية .

5) الصورة النهائية :

يمدنا تطبيق وتحليل الصورة الأولية ببيانات تفيدنا عند إعداد الصورة النهائية وهي الصورة التي تطبق في عمليات الثبات المرغوبة

التمييز بين الصواب والخطأ من خلال مجموعة عبارات ثم يضع أمامها علامة (√)أو (X) ولها بعض الشروط :

الا ترتب عبارات الخطأ ، الصواب في ورقة الأسئلة بطريقة معينة .

أن تكون العبارة مختصرة أو تشتمل علي نموذج للإجابة الصحيحة ، وعلي توجيهات واضحة ومختصرة .

ثانياً الاختيار من متعدد :

يقوم هذا النوع علي فكرة الاختبارات من أجل قياس قدرة التلميذ علي اختيار الإجابة الصحيحة لسؤال أو مشكلة من بين مجموعة من الإجابات التي تعرض عليه . وعلي هذا فالسؤال يتكون من جزاين .

ثالثاً : أن تكون من حيث الطول والاتزان بحيث تشغل أمهر التلاميذ طيلة وقت الامتحان وعلي الرغم مما يوجه إلي الامتحانات من نقد فأنه لا يمكن الاستغناء عنها . وذلك لأهميتها ولأنها أيضاً لا تخلو من فوائد .

فوائد الاختبارات الحديثة :

. أنها تزكي روح المهمة في التلاميذ وتمكن المثير منهم من بذل الجهد

. لها أثر تهذيبي عظيم في الحياة المدرسية .

. اهميتها عظيمة للمدرس فهي تقيس كفاءته كمدرس وتذكر في نفسه الحماس كي يتقدم ويعدل من طرق .

ـ أن الامتحانات تمدنا بمعرفة حالة كل طفل علي حدة ...... كما انها تقيس سر طرق التدريس أن كانت قوية أو ضعيفة .

. تشخص الصعوبات التي يلاقيها الطالب في دراسة المادة .

ويأخذ الامتحان شكلاً من أشكال ثلاث . سواء كان في نهاية كل فترة دراسية أم في نهاية العام أم في نصف السنة وذلك حسب الأمور الإدارية وكذلك المناهج ، إعدادها وأشكال الامتحان الثلاث هي :

1) الامتحان التحريري :

يكشف عن قدرة الممتحن علي ترتيب الأفكار وتنظميها كما يبرز أسلوبه ومدي هضمه للحقائق .

2) الامتحان الشفوي :

يكشف عن لباقة التلميذ وسرعة فهمه ودقة ذاكرته وحسن معالجته للمواقف الجديدة .

3) الامتحان العملي :

يكشف عن حروف أصابع التلميذ ومهاراته اليدوية وملاحظته وذكاءه العلمي .

مميزات الاختيار من متعدد :

من أجود أنواع الاختبارات وأكثرها مرونة . صياغته بأساليب مختلفة يصلح لقياس قدرة التلميذ علي عمليات عقلية راقية . فضلاً عن قياس التحصيل . ويحتاج هذا النوع لمهارة خاصة في وضعة .

شروطه :

- أن توضع البديلات فيه ذات أهمية خاصة .
- أن تكون البديلات في كل سؤال محصورة في دائرة ضيقة .
- أن توضع البديلات الصحيحة وغير الصحيحة في نظام غير محدد .

اختبار التكميل :

في هذا النوع يختار أو يضع الطالب الكلمات الناقصة خلال عدد من الجمل حتى يصبح المعني صحيحاً .

مميزاته :

تساعد التلميذ علي إبداء رأيه .

تغطي معظم الكتاب .

عيوبها :

ذاتية التصحيح

لا تصلح لقياس قدرات عليا مثل التحليل والابتكار .

شروطها :

أن يكون الفراغ المتروك في الجملة مكان كلمة أساسية ، إلا يكون هناك اختلاف علي الكلمة المطلوب وضعها في المكان الخالي .

يحسن عدم ترك مسافات كبيرة في الجملة الواحدة .

يحسن أن يكون طول المسافة المتروكة موحداً .

اختبار المزاوجة:

في هذا النوع من الاختبارات يعرض علي التلميذ قائمتان القائمة الأولي فيها عدد من العبارات . القائمة الثانية فيها إجابات هذه العبارات ولكن بترتيب مخالف .

ويجب أن تكون جميع العبارات متجانسة تدور حول موضوع واحد ، ويكون عدد الوحدات أو الفقرات بقائمة الإجابات أكبر من منها بقائمة الأسئلة .

ثانياً الاستخبارات :

الاستخبار أداة مهمة للبحث والقياس والاستخبار يتكون من مجموعة من الفقرات التي تعالج مشكلة أو قضية محددة أو جانب معين من جوانب الشخصية والتي يجيب عنها مجموعة من الأفراد بهدف جمع بيانات ومعلومات عنهم أو عن المشكلة أو القضية موضوع الاستخبار .

وللاستخبارات أنواع شتي تبعاً للجوانب التي نود معرفتها لدي المفحوص أو المفحوصين ، فهناك استخبارات للشخصية والاتجاهات والميول والقيم والدوافع والحاجات والجوانب الوجدانية والاجتماعية . ويطبق الاستخبار في موقف قياس

فردي أو جمعي ، ويجيب عنه المفحوص علي أساس معرفته بنفسه . مشاعره وانفعالاته وميوله وقيمه وحاجاته وسلوكه الماضي والحاضر ، وذلك بهدف التشخيص النفسي أو الإرشاد أو التوجيه المهني أو الاختيار المهني أو البحوث . وتصحح الإجابة بطريقة موضوعية دقيقة ، وتفسر الدرجة التي حصل عليها بالنسبة لمجموع التقنيين وقد يكون الاستخبار احادياً (يقيس جانباً أو سمة واحده كالميل إلي دراسة الميكانيكا أو القلق) أو متعدد الأبعاد (يقيس مجموعة من السمات أو جوانب السلوك) . ويوفر الاستخبار للمعلم بيانات ومعلومات صادقة وغير مباشرة عن تلاميذه في الجوانب الوجدانية والاجتماعية والتي يصعب التعبير عنها أو التصريح بها شفوياً أو كتابة .

ثالثاً : اختبارات الملاحظة :

يستطيع المعلم أن يجمع كثيراً من المعلومات العامة عن تلاميذه عن طريق الملاحظة المباشرة للتلاميذ في أثناء وجودهم في حجرة الدراسة أو عند قيامهم بأنواع النشاط المختلفة ويحسن أن يسجل المعلم ملاحظاته علي التلميذ بطريقة منظمة تبين سلوك التلميذ والتغيرات التي تطرأ عليه حتي يستطع المعلم أن يستخدم هذه الملاحظات في فهم سلوك التلميذ بمعني أن تكون الملاحظة دقيقة . وتوجد ملاحظة عابرة (غير مقصودة) ، وأخري مقصودة (علمية) وكلاً منهما يفيد في جمع البيانات واتخاذ القرار ، وتحتاج الملاحظة العلمية لأدوات وتسجيل البيانات أولاً بأول ، هذا بالإضافة إلي وجود هدف محدد لها .

وعلي الرغم من أن الملاحظة العلمية قد تستغرق وقتاً كبيراً إلا انها تمدنا بمعلومات قد يتعذر علينا معرفتها بدونها ومن النواحي الهامة التي ينبغي ان يهتم المعلم بملاحظتها في التلميذ :

. النواحي الاجتماعية والخلقية ، مثل التعاون . أساليب تعامل التلميذ مع الآخرين ، الأمانة ، الشجاعة .

- النواحي العقلية ، مثل الشغف بدراسة المواقف التي يشاهدها التلميذ . التفكير السليم الرغبة في البحث .

. النواحي الجسمية ، مثل ، الصحة الجيدة . الخلو من الإصابات والعاهات .

مشكلة الامتحانات :

هناك مجموعة من المعلمين يتخذون موقفاً عدائياً من الامتحانات ويؤكدون أنه لا مكان للقياس في التربية وذلك لان التربية تهتم أو تعالج أمور عقلية وأمور معنوية كالميول والآمال والمثل العليا .

بمعني أخر تهتم التربية بأمور هي أبعد ما تكون عن نواحي القياس والواقع خلاف ذلك . هؤلاء النفر يضطرون بطرق غير مباشرة إلي الامتحانات في صورة من صورها المختلفة كتشجيع التلاميذ . والإطلاع علي المجال الجيد والتفريق بين إجابات التلاميذ .

حتماً أن التربية تهدف إلي أحداث بعض التغيرات في المتعلم . تغيرات لا يمكن رؤيتها بالعين المجردة في حد ذاتها . لكن لابد من وسيلة تكون وظيفتها الأساسية إظهار هذا الجزء الحقيقي .

والامتحانات هي أعظم وسيلة لتحقيق هذا الغرض ، فعلي هؤلاء المعارضين للقياس في عملية التربية أن يسلموا بأن الامتحانات ضرورية . هذا ومعظم المعلمين يؤمنوا بأن الامتحانات في أي شكل جزء هام من العملية التربوية وانه من الصواب أن تصلح ما بها من عيوب بدلاً من نقضي عليها وإصلاح الامتحانات أمر يشغل بال رجال علم النفس التجريبي في الوقت الحاضر وهم يؤمنون بأن

امتحاناتنا المدرسية بصورتها الراهنة كثيرة المساوئ ويذكرون من هذه المساوئ ما يأتي :

1) ان هذه الامتحانات أصبحت هدفاً في حد ذاتها بدلاً من أن تكون وسلة لخدمة أهداف التربية . فالامتحانات اليوم تدفع بالقائمين علي شئون الأطفال إلي النصيحة بأغراض التربية وطرقها الحديثة . وأساليبها المختلفة فهم يقومون بأعداد التلاميذ وتهيئتهم للامتحانات بدلاً من تربيتهم الحقيقية التي تعدهم للكفاح في خضم الحياة فقد أصبحت مدارسنا في الوقت الحاضر معامل لتلقين النشئ نشوءا من المعرفة من المعرفة لارتباط بيئتها ، استفسار منها وأن كان لها أهميتها في الامتحان وبذلك فقدنا الكثير من الفوائد التي تعود علي النشئ لو حققت الامتحانات التربية الحقة .

2) ان الامتحانات بشكلها الحالي لا تساعدنا في الحكم علي الطفل حكماً حقيقياً . فهي لا تمثل إلا ناحية واحدة هي مبلغ افادته من المواد الدراسية أما عمق تفكيره ، عبقريته ، نشاطه ، وصلاحيته لحياة الاجتماعية فكل هذه اعتبارات أبعد ما تكون عن اهتمام الامتحانات الحالية .

3) ان الامتحانات بشكلها الحالي تشجع علي الاستظهار والحفظ وتقتل عمليات عقلية أرقي كالتفكير ، والقدرة علي الحكم ، فلا غرابة أذاً أن ينسي التلميذ كل شئ عقب انتهائه من الامتحان . ولذلك يلجأ إلي عم الملخصات . فالامتحانات الحالية لا تعرفنا شيئاً عما يعكسه التلميذ من تقدم في نواحيه الوجدانية ، الخلقية ، الاجتماعية . ولا تكشف لنا عن قدرته في استخدام ما إفادة من خبرة وتجارب من التصرف في المواقف الجديدة .

4) لقد أدي نظام الامتحانات إلي خطر صحي واخلاقي يتأثر به التلاميذ : الخطر الصحي يتمثل في أنه قبيل الامتحان تعتل صحة التلميذ ويرهقون

بالعمل والمذاكرة الطويلة مما يؤدي إلي إحداث أضرار جسمية ، ونفسية ، فقرب الامتحان نري وجوهاً شاحبة وأعصاباً مضطربة وقلق .... أما الخطر الأخلاقي .... ذلك في اقتران العلم دائماً بأهداف مادية فقد تدفع التلميذ إلي التحايل علي النجاح بأي شكل من الأشكال فقد يقتصر الاهتمام علي بعض أجزاء المنهج التي يتوقع منها الأسئلة كما يلجأ لي الغش بأي شكل من اشكاله وأيضاً يؤجل دراسته واستذكاره إلي أخر العام حيث يمكن أن يعتمد علي مدرس خصوصي يقوم بعملية تلقينه العلم ....   وهكذا .

5) الامتحانات الحالية قياس ذاتي يختلف باختلاف المصحح ويعوامل شخصية لا علاقة لها بمقدرة التلميذ مثل خط التلميذ وأسلوبه وطريقة عرضة وقدرته الإنشائية ، مزاج المصحح وفكرته عن التلميذ وعلي ذلك نعتقد أن الخط يدخل فيها تدخلاً كبيراً .

6) الاهتمام بالحصول علي درجات مرتفعه بالدرجة الأولي فهناك ما يشار إليه بحث الطالب علي الاهتمام بالنجاح الأولي ، وتأتي المعرفة ، والعلم في مرتبة ثانية .

7) أن عامل المصادفة ، عامل الحظ يلعبون دوراً كبيراً جداً فيهما ويتحكمان في تقدير مصير الطفل .... إذا أن هناك مواطن ضعف كثيرة في أسئلة الامتحان التقليدية .

فقد يتعرض التلميذ فيها للخروج عن موضع السؤال كما أن هذه الأسئلة كثيراً ما يشوبها عنصر الغموض فلا يفهم الطفل معناها .

8) يؤكد رجال التربية أن الأسئلة التقليدية تحتاج نضج عقلي قلما يصل إليه التلميذ حتي في مرحلة التعليم الثانوي . كما أن قلة عدد الأسئلة تؤدي إلي إغفال أجزاء كبيرة من المنهج وبذلك لا يستطيع قياس قدرة الطالب

فقد يصادف أن يعرف الطالب الكثير عن حقائق المادة ولسوء الحظ لا يمكن أن يجيب .

فقد ترتب علي ذلك نتائج غير سليمة منها :

. أصبح المدرس يعني بتدريسه للحقائق والمعلومات وكيف يتقنها التلاميذ بصرف النظر عن مدي أثرها في نفوسهم وبصرف النظر عن مدي قيمتها في حياتهم ، وأصبح المدرس أيضاً يوجه نظر التلاميذ باستمرار إلي مواضع المادة الدراسية التي تكثر فيها أسئلة الامتحانات .

. أحتج أولياء أمور التلاميذ بجنون أبنائهم علي تكديس الحقائق والمعلومات في اذهانهم عن طريق الاستذكار المتواصل الذي يجعل معظم عمل التلاميذ هو الاستظهار أكثر من دراسة منتظمة .

أصبح التلاميذ لا يتحمسون للقراءات المتنوعة أو للإطلاع الواسع المناسب لأنه يركز اهتمامه علي القراءات المحددة التي تطلبها الامتحانات ويشترط إتقانها كأساس للامتحانات .

. أصبح التلميذ في غمرة مطالب الامتحانات يهمل ما قد يكون لديه من قدرات خاصة و مهارات عملية معينة لا تفي أسئلة الامتحانات بها .

وتسعي الامتحانات الحالية لبلوغ الأهداف الزمنية :

. معرفة مدي فهم التلاميذ للمعلومات والحقائق .

. انتقال التلاميذ من صف دراسي إلي آخر .

. قبول التلاميذ في مرحلة معينة او تعليم معين .

. معرفة مقدار تحصيل التلميذ في نهاية مرحلة أو صف ومدي استحقاقه شهادة هذه المرحلة .

وعلي ذلك فان الامتحانات حينما تهمل الأهداف فانه يترتب علي ذلك عيوباً وساد اعتقاد مفاده أن امتحانات المقال هي أفضل الامتحانات وأن التقويم هو عملية تنتهي في نهاية المرحلة أو الصف الدراسي .

ومن الأمور التي تزيد من قدرة الامتحانات :

. لأن يغطي الامتحان جزءاً كبيراً من المنهج ويتناول موضوعاته الأساسية وأن لا تقتصر الأسئلة علي أبواب محددة .

. أن توضع الأسئلة بحيث تختبر مدي تقدم التلميذ نحو الأهداف المنشودة .

. أن توضع الأسئلة بحيث تميز مستويات التلاميذ وأن تحتوي علي التفكير والربط والاستنتاج.

. أن تكون هناك فرصة للاختيار لإظهار ميول التلاميذ .

. أن يعني الامتحان بالأحداث الجارية في البيئة لتوجيه اهتمام التلاميذ .

. عدم تشجيع استخدام كتب الملخصات بان تأتي الأسئلة محققة لها.

ويرجع اهتمام المعلمين بامتحانات المقال إلي كونهم تعودوا عليها في قياس تحصيل التلاميذ . حيث يلب فيه من التلميذ أن يستجيب استجابة حرة علي موضوع أو مشكلة .

من أجل ذلك اتجه رجال علم النفس التجريبي إلي البحث عن مقياس موضوعي لتقدير التلاميذ بدلاً من طريقة الامتحانات التقليدية .

فبحثوا عن مقياس يمكنه اختبار قوي التلميذ وقدراته العقلية ومقدار تحصيله في مختلف المواد . لأنه بدون مقياس موضوعي تصبح عملية القياس قليلة القيمة كما انها إذا ترك كل مدرس يقيس مدي نجاح تلاميذه بمقياس خاص تصبح النتائج التي يصل إليها عديمة القيمة كغيرها من المدرسين .

ولقد فكر رجال التربية في حل هذه المشكلة بطريقة فبدلاً من توضع في ورقة الامتحان خمسة أسئلة يحتاج كل منها إلي مقال للإجابة عنه ، فكروا في وضع عدد عظيم من الأسئلة قد يصل إلي المائة وكل سؤال فيه لا يحتاج في الإجابة إلي أكثر من بضع كلمات .

وطبق نفس المبدأ علي الرياضيات فيقدم لتلميذ عدد كبير في شكله العام شبيه باختبارات الذكاء ، ويطلق عليه أسم الاختبارات الحديثة .

الاختبارات المرجعية :

يمكن تقسيم الاختبار حسب الطريقة التي يفسر بها أداء التلميذ علي الاختبار إلي نوعين من الاختبارات ، الاختبارات مرجعية المعيار والاختبارات مرجعية المحك، والتميز أساسي بين النوعين هو أننا في النوع مرجعي المعيار نقارن أداء التلميذ بأداء الجماعة التي طبق عليها الاختبار أو بأداء الجماعة المعيارية ، في حين أننا في النوع مرجعي المحك نقارن أداء التلميذ بمستوي أداء معين أو محك محدد مسبقاً .

ولقد حدثت تطورات هامة في نظريات القياس التربوي والنفسي ونماذجه وتطبيقاته أثرت تأثيراً واضحاً في تطوير استراتيجيات التعليم وتصميماته ، وتحسين أساليب القياس وأدواته ، ومنهجيات التقويم وتقنياته . واهتم خبراء القياس والتقويم في العلوم التربوية والنفسية بتطوير الأساليب السيكومترية الكلاسيكية السائدة المتعلقة بالقياس مرجعي المعيار بما ينطوي عليه من فلسفة

تربوية تركز علي تصنيف الأفراد بحسب مركزهم النسبي بين أقرانهم في القدرات المختلفة ، وتوازن بين أدائهم في هذه القدرات دون تحديد لمكتباتهم الفعلية وكفايتهم الفردية .

وقد تبين محدودية هذه الأساليب الكلاسيكية وقلة جدواها في التشخيص الدقيق لهذه المكتبات والكفايات ، وكذلك قلة فاعليتها في إثراء المنظومة التعليمية وتطوير البرامج التدريبية ، إضافة علي ما يترتب علي نتائج الاختبارات والمقاييس التي تصمم وفقاً لهذه الأساليب من آثار اجتماعية ونفسية غير مرغوبة لدي الأفراد الذين تطبق عليهم هذه الاختبارات .

لذلك اتجه خبراء القياس و التقويم إلي تطوير مدخل أخر يستند إلي فلسفة تربوية منهجية مختلفة من أجل تلاقي أوجه القصور السابقة ، ويطلق عليه مدخل القياس مرجعي المحك ، ويركز هذا المخل علي قياس وتشخيص ما حققه المتعلم من معارف ومهارات ، وما أكتسبه من اتجاهات وسلوكيات تتعلق بمحتوي دراسي أو تعليمي أو تدريبي معين في ضوء محك أداء مطلق بغض النظر عن موازنة أداء المتعلم بأداء أقرانه فيما يقيسه الاختبار أو المقياس مرجعي المحك .

ولقد شهد مجال القياس مرجعي المحك تطورات ملحوظة سواء من الناحية النظرية أو من الناحية البنائية التطبيقية ونظراً حداثة مفهوم الاختبارات مرجعية المحك وتعدد تعريفها وتنوع استخداماتها ، فأننا سوف ندرس الاختبارات مرجعية المحك من حيث تعريفها ، الهدف منها ، أهميتها ، الأسس والسمات ، الاستخدامات ، أنماطها ، شروطها وبنائها وصدقها ثم ثباتها .

أولاً : تعريفها :

لقد أكد " جليزر " أن المقصود بالمحك هو أن نوازن أداء الفرد بنطاق شامل من المعارف والمهارات المحددة تحديداً جيداً كذلك أكد علي أن القياس محكي المرجع يعتمد علي مستوي أداء مطلق أي مستوي كفاية معين .

وقد عرفه " انور الشرقاوي "(1976) بانه ذلك الاختبار الذي يرد إلي ميزان والذي يذكر لنا شي عن الفرد بدون الرجوع إلي أداء فرد أخر . ففي المجال المعرفي مثلاً يذكر لنا الاختبار ماذا يعني الفرد أو ماذا يستطيع أن يفعل . وفي المجال غي المعرفي ، يكشف لنا عن الاتجاهات الخاصة بالفرد ويقيم مستوي أدائه.

وقد أكد "صلاح الدين علام " أن الاختبار محكي المرجع هو ذلك الاختبار الذي يستخدم لتقدير أداء الفرد في نطاق من السلوك محدداً تحديداً جيداً .

أي أن عبارة التحديد الجيد للنطاق السلوكي هي الركيزة الأساسية لهذا المفهوم ويعتمد تحديد النطاق السلوكي علي الأهداف أي المهارات المطلوب قياسها .

ثانياً : الهدف من الاختبارات مرجعية المحك :

يري " صلاح الدين علام " أن الهدف من الاختبارات مرجعية المحك هو اتخاذ قرارات تعليمية تختلف عن القرارات التي تتخذ في حالة الاختبارات جماعية المرجع . فاتخاذ القرار في هذه الحالة يكون علي أساس أداء كل فرد بالنسبة إلي مجموعة من المهام والمهارات والمعارف التي تشكل النطاق السلوكي الذي يقيسه الاختبار وهو بهذا يكون بمنزلة المحك الذي ينسب غليه الفرد .

ثالثاً: أهمية الاختبارات مرجعية المحك :

لاحظ " رالف تايلور" أنه اتضح من عدة دراسات لاختبارات التحصيل العالمية أن الأنواع الوحيدة في العناصر الاختبارية التي أدت إلي تباين كاف في

الاستجابة كانت العناصر التي تناولت محتوي المهارات التي لا يتم تعلمها في المدرسة . فلقد أصبحت اختبارات التحصيل عبارة عن اختبارات ذكاء . ولهذه الأسباب تضح اهمية الاختبارات مرجعية المحك لأنها سوف تصبح أكثر حساسية في تحيد أثار التعلم والاختبارات ذات الحساسية التعليمة لها قيمة كبيرة للقائم علي عملية التقويم .

رابعاً : الأسس التي يبني عليها الاختبار مرجعي المحك :

من الممكن تحديد المهارات والمعارف في محتوي معين والتي ينبغي علي الطالب إتقانها حتي يتقن المحتوي .

مفردة الاختبار مرجعي المحك يمكن أن تقيس بدقة مدي إتقان الطالب للمهارة التي تمثلها هذه المفردة ، وهذا يعني أن الطالب إذا أجاب علي المفردة إجابة صحيحة فقد أتقن المهارة أما إذا كانت الإجابة خطأ فهو لم يتقن المهارة .

خامساً : سمات الاختبارات مرجعية المحك :

اشار" مليمان "(1973) إلي سمات الاختبارات مرجعية المحك علي أساس أهميتها التربوية فهي تهتم بالعملية التعليمية ونواتجها وتساعد علي تعديل سلوك الطالب حتي بعد الوصول إلي المستوي المطلوب من الإتقان وأنها تستخدم باتجاه جديد في ان تحول اتجاهات الطلاب في التنافس علي الدرجات والتقديرات إلي التنافس في تحصيل الكفاءات والمهارات المطلوبة .

سادساً : استخدام الاختبارات مرجعية المحك :

- تستخدم الاختبارات المرجعية المحك في تشخيص العملية التعليمية المتتابعة وتقدم نمو التلاميذ وتعطي دلالة علي ما إذا كان التلميذ يحتاج إلي تقديم علاج اثناء عملية التعلم أم لا .

- تستخدم درجات الاختبار مرجعية المحك في تفسير النتائج تفسيراً مباشراً حيث أنها تتضمن حقائق وصفية دقيقة ومحددة عن ما يفعله الطلاب . كذلك تستخدم في اتخاذ القرارات التعليمية وتقويم البرامج وتأثيرها .

- يستخدم التقويم مرجعي المحك في تعديل المناهج وطرق التدريس بدرجة أكبر مما يمكن في حالة التقويم مرجعي المعيار .

- يمكن استخدامها في أجراء مقارنة بين تحصيل دفعات الطلاب الذين تخرجوا في فرقة دراسية معينة في أعوام مختلفة بدرجة كبيرة من الثبات .

. تمدنا بتعاريف إجرائية تعطي معني للدرجات التي يحصل عليها الطالب وكذلك لمستويات التحصيل علي أساس مطلق .

. تستخدم في تدريس المواد التي تتضمن تسلسلاً هرمياً من المهارات . حيث يعتبر تعلم بعض المهارات متطلبات أساسية سابقة لتعلم مهارات أخري . كما هو الحال في ميادين عديدة في العلوم والرياضيات .

. تشخيص مواطن القوة والضعف في تحصيل التلاميذ في مادة دراسية أو برنامج دراسي معين .

. تقويم اثر إجراءات تعديلات المنهج الدراسي في تحصيل التلاميذ .

. ضبط مدي تقدم التلاميذ بالنسبة للأهداف التربوية .

. تحديد مستوي التلاميذ بالنسبة للأهداف التربوية المقاسة .

سابعاً أنماط الاختبارات مرجعية المحك :

أشار " فؤاد ابو حطب " (1978) إلي وجود انماط متنوعة من الاختبارات مرجعية المحك أهمها : اختبارات إتقان ، واختبارات تشخيص .

1) اختبارات إتقان :

وهي تستخدم للتأكد من وصول التلاميذ إلي مستوي محدد من إتقان مادة معينة وهي غالباً ما تتصل بجزء صغير من المهارات أو الحقائق ولذلك فأن معظم الطلاب يسهل عليهم اجتياز هذه الاختبارات ويحققوا فيها من (90-100٪) إجابات صحيحة من مفردات الاختبار

وتشير " رمزية الغريب " (1970) غلي أن الغرض من اختبارات الإتقان هي تقييم التلاميذ إلي طائفتين أو نوعين هما : طائفة تعلم حقائق أساسية معينة . وطائفة لا تعرفها .

أما مستوي التلاميذ بعضهم بالنسبة لبعض فلا يعني واضع هذا النوع من الاختبارات **وأساسية التعلم للإتقان هي :**

. يختلف الوقت المسموح به للتعلم طبقاً لمعدل التعليم الفردي للتلاميذ .

ـ يحقق معظم التلاميذ مستوي الإتقان في متطلبات التعليم الفردي للتلاميذ .

ـ يحقق معظم التلاميذ مستوي الإتقان في متطلبات إذا سمح بالوقت الكافي للتعلم .

. يسجل مدي تقدم كل تلميذ ، باستخدام مراحل متعددة للتعلم الفردي .

. استخدام التقويم في كل مرحلة .

. يجب علي المدرس معرفة العناصر السلوكية والمهارات والمعارف ويحددها مسبقاً مع التلاميذ.

. تنافس التلاميذ يتم مع محك ، معيار وليس مع بعضهم .

وعرف " محمد عبد السلام " (1960) اختبارات الإتقان بأنه الاختبار الذي يكون غرضه الأساسي تحديد المدى الذي به تعلم أو تمكن الأفراد في جماعات في وحدة معينة من التدريب .

وهذا النوع من الاختبارات لا يقصد به التفريق بين الأفراد ولكن تحديد إذا ما كانت أو لم تكن مجموعة من الأفراد قد حصلت علي مستوي معين من الكفاءة .

2) اختبارات تشخيص :

ويري " فؤاد أبو حطب " أن اختبارات التشخيص تهدف إلي وصف الوضع الراهن للسلوك بهدف تحديد جوانب القوة والضعف والتعرف علي صعوبات التعلم ، كما يري أن الاختبارات التشخيصية تتشابه في كثير من خصائصها مع الاختبارات مرجعية المحك .

وأشار " محمد عبد السلام " إلي أنها اختبارات تستخدم في التشخيص وبيان نقاط الضعف ، والقوة في مجال معين من الدراسة وهو يتضمن مقاييس لمركبات أو أجزاء متفرعة في جهاز أكبر من المعلومات أو المهارات .

ثامناً : شروط الاختبارات مرجعية المحك :

اقترح " بابام " (1975 – 1978) تحديد النظام السلوكي الذي يقيسه الاختبار محكي المرجع باستخدام الخطوات أو المكونات الآتية :

1) وصف عام لما يقيسه الاختبار :

وهو عبارة عن وصف عام للسلوك فمثلاً الوصف العام في اختبار محكي المرجع يقيس قدرة الطالب علي التفكير بأسلوب علمي ويبقي ملاحظة أن الوصف

العام لما يقيسه الاختبار يمكن تعديله وتطويره بعد إجراء الخطوات التي تلي ذلك .

2) عينة من مفردات الاختبار :

وهي توضح لمن يقوم ببناء الاختبار كيفية قياس الهدف قبل البدء في قراءة تفاصيل المواصفات التي سيأتي ذكرها في الخطوات التالية .

3) تحديد عناصر المثيرات :

والمقصود بذلك تحديد ووصف المثيرات التي تشتمل عليها مفردة الاختبار . فالمفردة تعد بمثابة مثير يقدم إلي الطالب لكي يستجيب له بطريقة معينة وعناصر المثيرات تعتبر أهم مكونات مواصفات الاختبار إذا يعتمد عليها الفرد الذي يقوم ببناء الاختبار اعتمادا أساسياً .

ولذلك تختلف بالطبع درجة صعوبة عناصر المثيرات باختلاف المستوي الذي يشتمل عليه الاختبار . فكلما كان المحتوي عريضاً أو متسعاً مثل محتوي بعض المواد الدراسية كالتاريخ أو الأدب كلما زادت صعوبة هذه العناصر .

4) تحديد عناصر الاستجابة :

وتحديد ووصف طريقة استجابة الفرد لعناصر المثيرات التي تشتمل عليها مفردات الاختبار ويوجد نوعان رئيسيان من الاستجابات هما ، الاستجابات المقيدة (مثل الاستجابات لمفردات الاختيار من متعدد أو الصواب أو الخطأ أو المزاوجة) والاستجابات المفتوحة (مثل الاستجابات لمفردات المقال والمفردات التي تتطلب إجابات قصيرة) .

5) ملحق المواصفات :

ويضاف إلي ذلك ملحق المواصفات تكتب فيه التفصيلات التي تؤدي إلي مزيد من التوضيح لمحتوي النطق السلوكي أو محتوي مفردات الاختبار أو غير ذلك من عناصر المثير والاستجابة وهذا يقل من حجم التفاصيل التي تكتب في الخطوتين السابقتين وبذلك يمكن استخدام المواصفات بطريقة أكثر فعالية .

ثبات الاختبارات مرجعية المحك :

تعتمد طرق تقدير ثبات الاختبارات مرجعية المعيار علي مفهوم الارتباط بين متغيرين لذلك نجد أن قيم معامل الاستقرار . أو معامل التكافؤ تزداد بازدياد وتباين درجات الاختبار ويمكن أيضاً استخدام هذه الطرق في تقدير ثبات الاختبارات محكية المرجع إذا كان هناك قدراً معقولاً من التباين في الدرجات أما إذا كان توزيع درجات هذه الاختبارات متجانساً كما هو الحال في معظم الأحيان . وبخاصة إذا كانت الحالة تقدير ثبات الاختبار باستخدام أحد المقاييس اللابارامترية . وأن الاختبارات محكية المرجع تستخدم عادة في تصنيف الطلاب إلي مجموعتين أو أكثر بحسب درجة إتقانهم لأهداف محتوي دراسي معين لذلك أعتبر علماء القياس التربوي المعاصر أن ثبات هذا النوع من الاختبارات يعني أتساق التصنيف بناء علي الدرجات المستمدة منها . وتوصل عدد من هؤلاء العلماء إلي طرق مختلفة لتقدير الثبات لهذا المفهوم الجديد . وتحققوا من هذه الطريقة سواء باستخدام بيانات فعلية أو باستخدام أساليب المحاكاة عن طريق الحاسب الالكتروني مثل ذلك الأسلوب " مونت كارلو " وهو اسلوب يعتمد علي اصطناع بيانات تحدد فروض وتوزيعات البيانات العقلية .

ولا نقول بأنهم استطاعوا أن يتفقوا علي افضل هذه الطرق في تقدير الثبات ولا حسب في ذلك فالاختبارات محكية المرجع مازالت في مرحلة الطفولة . وتحتاج إلي المزيد من البحوث والدراسات.

صدق الاختبارات مرجعية المحك :

. الصدق: هو إجابة السؤال الذي يقول : ماذا يقيس الاختبار . وإذا لم يكن لدينا إجابة كافية لهذا السؤال فأن اي اختبار سوف يكون عديم الفائدة بالنسبة لما وضع لقياسه . وتختلف الاختبارات في مستويات صدقها تبعاً لاقترابها أو ابتعادها في تقدير تلك الصفة التي تهدف لقياسها .

ويتطلب ذلك تحديد العلاقة بين أداء المفحوص في الاختبارات ومجموعة من الحقائق والبيانات التي يمكن أن نلاحظها ملاحظة مستقلة عن الاختبار ونتناول الظواهر السلوكية موضوع الاهتمام . ويقاس الصدق للاختبارات مرجعية المحك بثلاث طرق :

1) الصدق الوصفي :

اختبار محكي المرجع صادقاً إذا استطعنا باستخدامه وصف أداء الفرد بالنسبة للنطاق السلوكي الذي يقيسه الاختبار .

والصدق الوصفي هو أول خطوة علي هذه الطريقة . وأحياناً يشار إلي هذا النوع من والسبب في اختيار مفهوم الصدق الوصفي هو أنه أكثر عمومية في صدق المحتوي وبخاصة إذا كان الاستخدام يقيس جوانب وجدانية أو نفسية حركية للسلوك حيث يكون مفهوم غير مناسب لهذه الجوانب . ويمكن تقدير هذا الصدق عن طريق جماعة من الخبراء والمحكمين في هذا المجال .

2) الصدق الوظيفي :

يقصد بالصدق الوظيفي أن يؤدي الاختبار محكي المرجع الوظيفة التي صمم من اجل تحقيقها أو الغرض الذي بني من اجله ولا يقتصر فقط علي الوصف .

وفي مرجع أخر فان الصدق الوظيفي هو الصدق الذي يطلق علي الاختبارات التي تستخدم لاتخاذ القرارات ذات الصبغة الوظيفية ووفاء الاختبار مرجعي المحك بتحقيق الغرض الذي وضع من أجله ودقته في ذلك .

3) صدق اختيار النطاق :

فالاختبار مرجعي المحك يكون صادقاً صدق اختيار النطاق عندما نستطيع تحديد نطاقاً معيناً اختاره واضع الاختبار يستخدم علي نحو دقيق وصحيح لبعد أكثر عمومية .

وفي مرجع أخر هو ذلك النوع الذي يناظر الصدق التجريبي والسبب في هذه التسمية هو أنها أكثر شمولية .

أوجه التشابه والاختلاف بين الاختبارات مرجعية المعيار ومرجعية المحك :

أولاً : أوجه التشابه :

تتشابه كل من الاختبارات مرجعية المعيار والاختبارات مرجعية المحك في بعض النقاط وهي كالأتي:

1. وجد أن كل من الاختبارات مرجعية المحك . والاختبارات مرجعية المعيار تتفق من حيث كونها اختبارات أداء تؤخذ علي شكل ورقة وقلم أو كنشاط يحتاج إلي تكملة الأعمال المحددة . فالاختبار الذي يجري للحصول علي رخصة القيادة يعتبر مثالاً لدمج كلاً من الامتحانات التحريرية وامتحانات الأداء معاً .

فالفرد المتقدم لاستخدام رخصة القيادة يجب عليه أن ينجح في الاختبارات التحريرية لتثبيت معرفته بقواعد المرور علي الطريق . ثم بعد ذلك يجب عليه ان

يقود السيارة بالفعل علي طريق مخصص للقيادة وذلك حتي يظهر إتقانه للمهارات المطلوبة في القيادة .

2. تتفق كل من الاختبارات التحصيلية مرجعية المحك والاختبارات مرجعية المعيار من حيث كونها اختبارات تحصيل أو اختبارات قدرة . والغرض من الاختبار التحصيلي هو إعطاء دلالة لما حصله التلميذ من قبل أما اختبار القدرة فالغرض منه إعطاء دلالة لما يتنبأ به وما يتوقعه من حيث تحصيل التلميذ في المستقبل . ونقلت " أمال محروس" عن "بيجز لويس " (1975) أن كلا النوعين متشابهان في :

أ- أن كل منهم أما أن يكون اختبارات تتعلق ببعض الأعمال التحريرية أو يكون اختبار ا\أداء بعض الأعمال محددة ز

ب- كلا النوعين اختبارات تحصيل أو استعدادات فالاختبارات التحصيلية تقيس ما يمكن أن يفعله الفرد في المستقبل .

ج- كلاهما يعطي معلومات عن مدي نجاح النظام التعليمي برغم أن كلاً منهما له دوره في تقويم النظام التعليمي .

أيضاً نقلت " امال محروس " عن " كرامبرت " (1975) عن أوجه التشابه بين الاختبارات المرجعية المعيار والمرجعية المحك :

1. من حيث التطبيق:

إن كلا النوعين من القياس يمكن الإسهام في وضع القرارات التعليمية إلا انه من الأفضل استخدام القياس مرجعي المحك في حالة القرارات التعليمية قصيرة المدى والقياس مرجعي المعيار في حالة القرارات التعليمية طوية المدى .

2. الثبات :

اشارات الدراسة إلي أن معامل كيودر ريتشارد سون (الصيغة 5ج) هي الملائمة لإيجاد الاتساق الداخلي للاختبارات مرجعية المحك . وأكدت ان الخطأ المعياري للقياس له أهمية بالنسبة للاختبارات مرجعية المحك كأهميته بالنسبة للاختبارات مرجعية المعيار.

3. الصدق :

أشارت الدراسة إلي استخدام الطرق التقليدية لتقدير صدق المحتوي ، وصدق المحك في الاختبارات مرجعية المحك كالطرق التي تتبع في الاختبارات مرجعية المعيار .

4. إمكانية الاستخدام :

أشارت الدراسة إلي أنه لا يوجد اختلافات جذرية بين الاختبارات مرجعية المحك ومرجعية المعيار يمكن من التمييز بينهما باستثناء درجة القطع ولكن الاختلاف في استخدام كل منهما .

ثانياً : أوجه الاختلاف :

لقد كان مستوي أداء الطالب يحدد لسنوات عديدة علي أساس نسبي مرجعي المعيار وكما نري فان درجه الطالب تحدد علي أسس متفقة بالنسبة لزملائه وليس علي أساس مدي قيمة الأداء بالنسبة لأي محك مطلق ولعل ممارسة مثل هذه التقديرات داخل الفصل لها تأثيرات غير مرغوب فيها وهما :

. أعطيت تأكيد للراي القائل انه لكي يكون النجاح له معني فانه يجب ان يكون هنالك مجموعة من الراسبين وعلي هذا الأساس يؤكد هذا النظام علي النجاح والرسوب بصرف النظر عن اكتساب التلميذ للمهارات التعليمية المطلوبة .

- أدي استخدام المقياس مرجعي المعيار إلي الاختلاف بين نظام الإثابة أو بمعني التقديرات وبين الأداء الفعلي للتلميذ بمعني الإثابة علي أساس تحصيلهم بالنسبة لزملائهم وليس علي أساس تحصيلهم بالنسبة إلي محك مطلق للأداء .

- اما الاتجاه مرجعي المحك يوجه الانتباه إلي نواحي هامة في العملية التعليمية وهي المهارات المحكية فهو يخدم وظيفتين هامتين داخل المضمون التعليمي هما :

- يوجه الانتباه إلي السلوك والإنجازات وهذا هو الهدف الأساسي من التعليم فيجب علي المدرس أن يهتم بهذا السلوك المحك وهل أتقن الطالب هذا السلوك ولا يكون اهتمامه منصباً علي مقدار ما اتخذه بالنسبة لزملائه .

ونقل " سليمان محمد سليمان " (1986) عن " جروتلند " انه يقلل من سعة الفجوة بين القياسات مرجعية المحك ومرجعية المعيار فيقول ..... من بعض الحالات قد يكون من المرغوب فيه إدماج نتائج التقويم المرجعي المعيار، والمرجعي المحك معا في نظام ثنائي يشتمل ذلك النظام علي نوعين من البيانات أحدهما: يحدد أداء الطالب بالنسبة لزملائه ، والأخر يحدد الدرجة التي بها حقق الطالب أهداف المنهج الذي درسه .

إذا كان المقياس مرجعي المعيار يستخدم لتقدير أداء الفرد بالنسبة إلي أداء الأفراد الآخرين من نفس مجموعته في القدرة التي يقيسها الاختبار فان القياس مرجعي المحك يستخدم لتقدير أداء الفرد بالنسبة إلي محك او مستوي أداء مطلق دون الحاجة إلي موازنة أداءه بأداء الأفراد الآخرين فالاختبار مرجعي المحك يصمم أساساً للإجابة علي سؤال هام " ماذا يستطيع الطالب أن يؤدي .... ؟ وماذا لم يستطيع أداءه ... ؟ أو ماذا أتقن ؟، وماذا لم يتقن ..... ؟ وهو بذلك

يفيد في تشخيص مواطن القوة والضعف في تحصيل الطلاب وهو ما يهم المعلم بالدرجة الأولي . ويفيد في توجيه عمليتي التعليم والتعلم .

نقلت " امال محروس " عن " بيجز ولويس "(1975) ان الرفق بين القياس مرجعي المعيار والقياس مرجعي المحك قد ينشأ عن اختلافهما في الهدف حيث الاختبارات مرجعية لتظهر اي أفراد المجموعة يكون أكثر قدر من المعرفة عن المحتوي المراد قياسه . وهذه الاختبارات تضم كي يتم توزيع الدرجات فيها علي أساس أن الدرجة الكبرى تشير إلي الأداء المفضل ، والدرجة الأقل تشير إلي الأداء الأدنى .

اما الاختبار مرجعي المحك فيحاول تقدير السلوك الذي تم قياسه وليس هناك محاولة للتمييز بين الأفراد فالجميع ربما يحصلون علي درجة المفردة الصحيحة . وأن التمييز بين المفردات الجيد والرديئة للاختبار مرجعي المحك يقوم علي أساس قياس المفردة للسلوك المراد قياسه أو عدم قياسها له .

أوضح " بايام " (1978) أن الفرق الرئيسي بين الاختبارات مرجعية المحك ومرجعية المعيار هو أن الأخيرة تركز علي تحديد المكانة النسبية للتلميذ بالنسبة للتلاميذ الآخرين وفيها يرجع الأداء الكلي للتلميذ إلي معيار او مجموعة معيارية .

بينما في الاختبارات مرجعية المحك يكون التركيز علي المستوي المطلق للعينة ، وعلي تحديد الوضع النهائي لأداء الممتحن اي أن درجة التلميذ تعطي معلومات واضحة عما يمكن أن يؤديه ، وفي هذا المقياس يرجع أداء التلميذ إلي محك بمعني مجال محدد جيداً من السلوكيات ويحدد أداء التلميذ في ضوء هذا المستوي الكلي .

الفصل الثالث

" شروط القياس النفسي "

أصبحت أهداف التربية غير مقصورة علي ناحية واحدة من نواحي تكوين الشخصية وهي ناحية تحصيل المعارف والأدوات الأساسية . وإنما امتدت لتشمل جميع مظاهر شخصية الإنسان العقلية والمزاجية والاجتماعية .... الخ ، لتعمل علي نموها ومدها بما يساعد علي الاستفادة من كافة إمكانيتها في الاتجاه المرغوب فيه .

ولذلك امتدت أدوات القياس بدورها لتشمل جميع هذه المظاهر إلا انه يجب أن نفرق بين مجموعتين من أدوات القياس . الأولي من عمل غير المدرس ، والثانية يقوم بها المدرس أو الهيئات التعليمية المباشرة .

وتحت المجموعة الأولي تأتي اختبارات الذكاء والاستعدادات الخاصة ، واختبارات الشخصية ، وكذلك اختبارات التحصيل التي توضع لأغراض تجريبية أ لأغراض التشخيص أو للاستفادة منها علي نطاق عام .

أما المجموعة الثانية فتأتي تحتها الاختبارات التحصيلية التي نستخدمها لأغراض مدرسية محددة .

وأياً كان نوع الاختبار النفسي ، فإن هناك شروطاً أساسية يجب مراعاتها لضمان سلامته ونجاحه وتحقيقه للأغراض المطلوبة منه ، فالاختبار أو القياس يجب ان يكون موضوعياً لا يتأثر بالعوامل الذاتية او بأي عوامل غير ذات صلة بموضوعه سواء في وضعه أو تصميمه .

والمقياس يجب أن يكون ثابتاً بمعني ان نفس النتائج عندما تطبق في الظروف المتشابهة وعلي نفس العينة من الأفراد ، أو نتائج متقاربة بدرجة مقبولة

(إحصائياً) وان يكون صادقاً يقيس ظاهرة محددة هي التي وضع من اجل قياسها وان تكون درجاته ذات معني ودلالة ، ويمكن ان نميز بين الأفراد أو بمعني أخر ان يكون له معيار او معايير تحدد درجة اداء الفرد بالنسبة لغيره من الأفراد أو بالنسبة للمجتمع الذي ينتمي إليه . فضلاً عن الشروط الاخري الخاصة باستخدامه وسهولة تطبيقه وتصميمه حتي يمكن الاستفادة منه كأداة ميسرة غير معقدة لتحقيق الأغراض المطلوب منه تحقيقها .

وهكذا نصل إلي ان الشروط الأساسية التي يجب توافرها في المقاييس النفسية هي :

1. الموضوعية .
2. الثبات .
3. الصدق .
4. ان يكون لها معيار او معايير محددة .
5. سهولة الاستخدام .

وهذه الشروط تراعي عادة في البحوث السيكولوجية وعند وضع الاختبارات التي تخدم أغراضاً عامة ، أو عن تقنين هذه الاختبارات ، من ثم فهي تحظي بالاهتمام بالنسبة للمجموعة الأولي من الاختبارات (اختبارات القدرات العقلية والشخصية ..... الخ)وان كان من الضروري بالنسبة للمجموعة الثانية من الاختبارات (اختبارات التحصيل المدرسية العادية) ان يضع المدرس في اعتباره مفاهيمها العامة ، وأن يعمل علي ضوئها اثناء وضع اختباراته أو تصميمها ، ليضمن موضوعيتها وسلامة الغرض الذي تؤديه وليضمن ايضاً سلامة النتائج المستمدة منها وتفسيرها الصحيح .

أولاً : الموضوعية Objectivity .

موضوعية الاختبار شرط أساسي ، والاختبار الموضوعي هو ذلك الذي لا يتأثر الشخصي لواضع الاختبار أو لمصلحته أو بغير ذلك من المؤثرات وتعتمد نتائجه علي الحقائق المتعلقة بموضوع الاختبار وحده ولذلك تصاغ أسئلة الاختبارات الموضوعية عادة في صورة كلمات أو عبارات أو أرقام أو أشكال .... الخ ، يجاب عنها إجابة محددة لا يختلف بشأنها ، قد تكون علامة توضع تحت كلمة (صحيح) أو (خطأ) أو تكون الإجابة كلمة بالذات تترك لها مسافة في ورقة الإجابة ، أو قد تكون رقماً معيناً ..... أو نحو ذلك مما لا يدع مجالاً لتأثير الشخصي أو الذاتي للمصحح ،بحيث تكون درجة الشخص في الاختبار واحدة مهما تغير المصححون .

وهناك حالات قليلة تفضل فيها الاختبارات غير الموضوعية وأغلبها ذات طابع إكلينيكي يعتمد تقدير النتائج فيها علي الانطباع الشخصي والرأي الذاتي للأخصائي النفسي . ولذلك يجب أن تكون من المرونة بحيث تتيح إمكانيات أوسع للتقدير والتفسير وأن كان من الأفضل وبالرغم من هذه الأسباب . أن توضع لها ضوابط موضوعية حتي لا تترك تماماً للتقديرات الذاتية .

ولضمان موضوعية الاختبار يجب أن يخضع وضعه وتصميمه والظروف التي يجري فيها وغير ذلك من المؤثرات والمتغيرات الخاصة بالاختبار لشروط موحدة ، حتي تظل هذه العوامل ثابتة وحتى لا يكون هناك متغير عن الفرد نفسه الذي يجري عليه الاختبار ، وحتى تكون الدرجة المعطاة متوقعه عليه وحده ولا علاقة لها بأي عامل أو مؤثر اخر .

ولتحقيق هذه الغاية يجب أن ننتبه لمجموعة من الأشياء ، فمثلاً يجب أن تكون الأسئلة واضحة تماماً حتي لا يتدخل عامل عدم فهم العبارات أو الأسلوب

المستخدم ، واختلاف الأفراد بعضهم عن بعض في هذه النواحي في درجة الاختبار، وان تكون صياغة الأسئلة دقيقة بحيث لا يختلف الأفراد في تفسيرها ومعرفة المقصود منها ، إلا تشتت إجاباتهم تبعاً للتفسيرات المختلفة التي توحي بها ، والتي ليس لها علاقة بالموضع نفسه الذي يقيسه الاختبار .

وخير طريقة لتجنب هذه الأخطاء هي تجريب الاختبار قبل استخدامه علي عينة من الأفراد مماثلة للأفراد الذين يوضع من أجلهم الاختبار لتحديد الأسئلة غير المفهومة أو التي لها أكثر من تفسير وحذفها أو تعديلها .

ويحسن أيضاً أن يعرف الفرد مقدماً الطريقة التي سيصحح بها الاختبار حتي يكون التقدير المعطي له سليم فإذا كنا سنحاسبه على الإجابات الصحيحة وحدها فيجب أن يعلك ذلك أيضاً .

وفي جميع الأحوال يجب أن يقدم الاختبار بتعليمات وتوضح طريقة العمل وطريقة الإجابة تماماً ، ويجب أن تضمن التعليمات مثلاً أو أكثر ليعرف الفرد المطلوب منه ويتجنب الأخطاء التي تأتي نتيجة عدم فهمه أو اهتمامه لها ولا تتداخل هذه العوامل في أداء الفرد وتتأثر بها درجته في الاختبار وحتى يقتصر أداؤه علي الحقائق أو النواحي التي هي موضوع الاختبار .

ثانياً : الثبات Reliability .

أن ثبات الاختبار هو أن يعطي الاختبار نفس النتائج إذا ما أعيد علي نفس الأفراد في نفس الظروف أي أنه يعطي نفس النتائج إذا قاس نفس الشئ مرات متتالية ، والثبات في علم القياس النفسي هو دقة الاختبار في القياس أو الملاحظة وعم تناقضه مع نفسه واتساقه وإطراده فيما يزودنا به من معلومات عن سلوك المفحوص .

طرق حساب الثبات :

توجد طرق مختلفة لحساب الثبات ، وهذه الطرق ليست بدائل يحل بعضها محل بعض وإنما لكل منها استخدامه الخاص حسب تصور تباين الخطأ الذي نهتم به ومن هذه الطرق ما يلي :

1) إعادة الاختبار :

تقوم فكرة هذه الطريقة علي إجراء الاختبار علي مجموعة من الأفراد ثم إعادة إجراء نفس الاختبار علي نفس مجموعة الأفراد بعد مضي فترة زمنية ، وهكذا يحصل كل فرد علي درجة في الإجراء الأول للاختبار ، وعلي درجة أخري في الإجراء الثاني للاختبار وبحساب معامل الارتباط بين الدرجات في المرة الأولي والدرجات في المرة الثانية فإننا نحصل بعد ذلك علي معامل ثبات الاختبار.

ويستخدم معامل الثبات وهو نوع من معامل الارتباط للتعبير عن مقدار ثبات درجات الأفراد في نفس الاختبار . إذا طبق عليهم مرتين بينهما فاصل زمني مناسب (ينحصر بين أسبوعين وستة أشهر) .ويعبر معامل الارتباط عن مدي التماثل أو العلاقة بين مجموعتين من الدرجات .

والقيم التي نحصل عليها من معامل الارتباط تمتد من (- 1.0 إلي +1.0) ونادراً ما نحصل علي معامل ارتباط (ر=- 1.0) سالب تام ، أو (ر=+1.0) موجب تام . وعلي سبيل المثال إذا كان معامل الارتباط بين الذكاء والتحصيل مقداره (ر=0.80)فهذا يعني أن 80٪ من عدد أفراد العينة يتفق مستوي ذكائهم مع مستوي تحصيلهم ، بينما 20٪ من أفراد العينة لا يتفق مستوي ذكائهم مع مستوي تحصيلهم .

ويحسب معامل الارتباط بين درجات أفراد العينة بين متغيرين بعدد من الطرق منها طريقة " بيرسون " ، وذلك باستخدام الدرجات الخام .

معامل الارتباط =

$$\frac{\text{ن مجـ س ص} - \text{مجـ س} \times \text{مجـ ص}}{\sqrt{\{\text{ن مجـ س}^2 - (\text{مجـ س})^2\}\ \{\text{ن مجـ ص}^2 - (\text{مجـ ص})^2\}}}$$

باستخدام متوسط الانحرافات

$$\text{معامل الارتباط} = \frac{\text{مجـ ح}_1\text{ح}_2}{\text{ن} \times \text{ع}_1\text{ع}_2}$$

باستخدام معامل ارتباط الرتب :

$$\text{ر}_{\text{س ص}} = 1 - \frac{6\ \text{مجـ ف}^2}{\text{ن}(\text{ن}^2 - 1)}$$

حيث يدل الرمز ر س ص علي معامل ارتباط الرتب ، ويدل الرمز مجـ ف^2 علي مجموع مربعات فروق الرتب ، والرمز ن علي عدد الأفراد .

مثال : احسب معامل الارتباط بين درجات مجموعة من الأفراد عددها (12) في مادتي الرياضيات والعلوم من الدول التالي :

م	الرياضيات	العلوم	م	الرياضيات	العلوم
1	30	35	7	45	40
2	37	32	8	42	45
3	40	41	9	28	30
4	20	30	10	33	35
5	25	30	11	25	36
6	35	30	12	24	30

وتنتقد هذه الطريقة في تدخل عدد من العوامل بين التطبيقين للاختبار منها الممارسة أو الخبرة بالاختبار التي قد تنتج قدراً من التحسن في درجات الاختبار عند إعادة تطبيقه ومنها أيضاً أنه لو كانت الفترة بين موقف الاختبار وإعادته قصيرة فأن المفحوص قد يستدعي بعض استجاباته السابقة ، ومنها كذلك أن طبيعة الاختبار ذاتها تتغير بسبب الإعادة وخاصة إذا كانت أسئلة الاختبار تتطلب استخدام بعض العمليات التي تحدث في موقف تطبيق الاختبار نفسه والتي قد يكون مسئول عنها الفاحص أو ظروف المكان أو تغير وقت تطبيق الاختبار ، ويقترح أن تكون الفترة الزمنية بين التطبيقين اسبوعين وستة أشهر ، ويحدد مضمون الاختبار هذه الفترة الزمنية كمال تحددها طبيعة الاختبار نفسه .

2) تنصيف الاختبار " التجزئة النصفية ":

في هذه الطريقة نطبق الاختبار علي العينة مرة واحدة ، ثم يقسم إلي نصفين متساويين ، ويحسب لكل فرد في العينة درجتاه في النصفين

وتوجد عدة طرق لتنصيف الاختبار وفي الطريقة الأولي نقوم بقسمة عدد مفردات الاختبار إلي نصفين متساويين ، فإذا كان الاختبار يتكون من (50) مفردة فإن النصف الأول سيكون عبارة عن البنود من (1 إلي 25) ، والنصف الثاني من (26 إلي 50) . النصف الأول من المفردات أرقام 1، 3، 5، 7 ، .... ، والنصف الثاني من المفردات أرقام 2، 4 ، 6 ، 8 ،.........، وهذه الطريقة . التنصيف الفردي والزوجي . تتلافى عيوب الطريقة السابقة . التنصيف إلي نصفين متساويين . حيث تؤدي غلي ضبط متغير دخيل مثل عوامل التعب والإرهاق والملل حيث تؤثر في الجزء الأخير من الاختبار بصورة سالبة .

وتوجد عدة معادلات لحساب معامل الثبات بالتجزئة النصفية منها:

1- معادلة سبيرمان وبراون :

$$\text{ر}_{11} = \frac{2\text{ر}}{1 + \text{ر}}$$

حيث يدل ر11 علي معامل الثبات ، ر علي معامل الارتباط بين درجات النصف الفردي والنصف الزوجي باستخدام معادلة بيرسون . ولا تصلح هذه الطريقة إذا انقسم الاختبار لقسمين غير متكافئين ، حيث يفترض هنا تكافؤ الجزأين . ولا تصلح أيضا لحساب ثبات الاختبارات الموقوتة التي يترك فيها الأفراد عدد من الأسئلة بدون إجابة .

2- معادلة رولون :

$$\text{ر}_{11} = 1 - \frac{\text{ع}^2\text{د}}{\text{ع}^2}$$

حيث يدل الرمز $ع^2$ف على تباين فروق درجات النصفين ، $ع^2$ على تباين درجات الاختبار ككل .

3- معادلة جتمان للتجزئة النصفية :

$$ر_{11} = \frac{2\left(1 - (ع_1^2 + ع_2^2)\right)}{ع^2}$$

حيث يدل الرمز $ع_1^2$ على تباين درجات الأسئلة الفردية ، $ع_2^2$ على تباين درجات الأسئلة الزوجية ، $ع^2$ على تباين درجات الاختبار .

وتستخدم هذه المعادلة عندما لا تتساوي الانحرافات المعيارية لجزئي الاختبار كما تستخدم أيضاً في حالة تساوي هذه الانحرافات .

4- معادلة الفا كرونباك :

$$ر_{11} = \frac{ن}{ن - 1}\left\{1 - \frac{مجـ ع^2 جـ}{ع^2}\right\}$$

حيث يدل الرمز ن على عدد المفردات ، ويدل الرمز مجـ ع 2 جـ على مجموع تباين المفردات لدي كل الأفراد ، ع 2 على تباين الدرجة الكلية لدي كل الأفراد ، وتصلح هذه الطريقة (ألفا . كرونباك) لكل أنواع الاختبارات سواء التي فيها نظام التصحيح (صفر ، 1) أو متدرج .

5- معادلة كيودر. ريتشارد سون (KR 20)

$$ر_{11} = \frac{ن}{ن - 1}\left\{1 - \frac{مجـ س ص}{ع^2}\right\}$$

حيث يدل الرمز ن علي عدد الأسئلة ، الرمز س علي معامل سهولة السؤال ، الرمز ص علي معامل صعوبة السؤال ، ع2 تباين درجات الاختبار .

ويشترط لتطبيق هذه المعادلة ان تكون درجة اي سؤال هي واحد أو صفر . أي تعني الإجابة الصحيحة واحد ، والإجابة الخطأ صفر .

6- معادلة كيودر. ريتشارد سون (K R 21):

$$\text{ر}_{11} = \frac{\text{ن}}{\text{ن} - 1}\left\{1 - \frac{\text{م}(\text{ن}-\text{م})}{\text{ن}\,\text{ع}^2}\right\}$$

حيث م متوسط درجات الأفراد في الاختبار ، ن عدد الأسئلة وتصلح هذه المعادلة في حالة الاختبار التي تتدرج فيها الدرجة .

2) الصور المتكافئة :

وتقوم هذه الطريقة علي فرض أساس هو أنه من الممكن اعداد صورتين متكافئتين من الاختبار الواحد . وفي هذه الحالة يكون معامل الارتباط بين نتائج تطبيق الصورتين المتكافئتين هو معامل الثبات علي نفس الأساس كما لو كان معامل الارتباط بين نتائج الاختبار ونفسه . وتكمن صعوبة هذه الطريقة في بناء صورتين متكافئتين ، وكذلك فإنه الألفة بموقف التطبيق للصورة الأولي قد يؤثر علي الأداء في الصور الثانية إلا أن تذكر الإجابات يكون اقل لأن المفردات ليست هي نفسها .

وتوجد طريقتين لاستخدام الطرق المتكافئة الأولي، نستخدم فيها صورة مكافئة فورية يتم تقديمها في موقف الاختبار نفسه تالية للصورة الأولي للاختبار . وهناك لا يوجد تقدير للتباين في النتائج عن العوامل الزمنية أو استقرار الأداء علي مدي زمني . بل تقدير لاستقرار الأداء من خلال الاتساق في عينة الاختبار

بين الصورتين المتكافئتين ، والطريقة الثانية نستخدم فيها صوراً متكافئة متعاقبة وتقوم بعد فترة زمنية لاحقة ،وهنا يتعين معامل الثبات الناتج تقديراً لكل من الاتساق في عينة مادة الاختبار والاتساق في الأداء علي مدي الزمن ،ويجب ذكر طول الفترة المنقضية بين استخدام الصورة الأولي والصورة الثانية .

العوامل التي تؤثر علي ثبات الاختبار :

1) عدد مفردات الاختبار " طول الاختبار ": كلما زاد عدد مفردات الاختبار كلما أرتفع ثباته .

2) زمن الاختبار :

يزداد ثبات الاختبار تبعاً لزيادة الزمن حتي يصل إلي الحد المناسب ، وعند ذلك يصل الثبات إلي نهايته العظمي ثم يقل بعد ذلك كلما زاد الزمن عن ذلك الحد .

3) التباين :

يرتبط الثبات بتباين درجات الاختبار ، حيث يزداد ثبات الاختبار تبعاً لزيادة التباين .

4) صعوبة المفردات :

تؤثر صعوبة مفردات الاختبار علي ثباته ، فكلما كانت مفردات الاختبار اكثر تمييز كلما أرتفع ثباته.

5) التخمين :

إذا نتج عن التخمين إجابة صحيحة فإنه يسهم بمقدار من الخطأ في درجة الفرد في الاختبار ومقدار من الخطأ في تفسير هذه الدرجة .

6) الفترة الزمنية بين التطبيقين للاختبار :

إذا كانت المدة الزمنية طويلة بين التطبيقين فقد تتغير القدرات الحقيقة للفرد ، وإما إذا كانت المدة قصيرة فان معامل التذكر يؤثر. لذا يجب تحديد الفاصل الزمني بين التطبيقين في طريقة إعادة الاختبار والصور المتكافئة .

ثالثاً : صدق الاختبار : Validity

يتعلق صدق الاختبار بما يقيسه ، وكيف يقيس ، ويخبرنا بإمكانية الاستدلال من درجات الاختبار ، وهذا الربط أو التسلسل المنطقي يجعلنا في وضع دفاع ضد القبول أو التسليم بأن اسم الاختبار دليل لما يقيسه . إن اسم الاختبار يمدنا باختصار بوصف ملائم للتعرف علي الأهداف ، وغالباً ما يكون أسم الاختبار ابعد بكثير عن الغموض والوضوح ، يزودنا بدلالة المعاني بتغطية مجال السلوك بالرغم من المجهود الزائد المبذول لاستخدام نوعية أكثر وأسماء اختبارات معروفة امبريقياً . أن السمة المقاسة بواسطة الاختبار من الممكن تحديدها فقط من خلال امتحان في المصادر الموضوعية للمعلومات والفائدة العلمية لإظهار صدقها ، وفوق ذلك فإن صدق الاختبار لا يمكن أن يقرر أو يحدد من عبارات عامة ، ولا يمكن في مقالة أن نقول إن الاختبار مرتفع أو منخفض الصدق . أن الصدق يجب تعيينه من خلال الرجوع إلي الاستخدام الخاص للاختبار والذي من خلاله يكون مدروساً . علاوة علي ذلك ، فإن جميع طرق إيجاد صدق الاختبار تتعلق بالعلاقة بين الأداء علي الاختبار وعوامل ملاحظة أخري مستقلة عن صفات السلوك المدروس ، أن الطرق الخاصة المستخدمة لبحث هذه العلاقة عديدة وعرفت بمسميات مختلفة ، وأن هذه التوضيحات والأوصاف التقليدية ركزت علي مظاهر مختلفة للصدق ، ونظراً للاهتمام الخاص بالاستخدامات المختلفة

للاختبارات وتطور واتساع تطبيقها ، فإن مفاهيم الصدق تغيرت وعدلت علي نحو مطابق .

تطوير مفهوم صدق الاختبار :

من بين أقدم استخدامات الاختبارات ، استخدمها في تقدير ما تعلمه الأفراد في مجالات ومناهج محددة ، والآن يتمثل ذلك الاستخدام في امتحانات نهاية المقرر الدراسي ، وفي اختبارات الترخيص لقيادة سيارة أو للتأهيل لمهنة معينة . وهذا النوع من الاختبار ، الذي يصمم كاختبار للتحصيل ، يقيم بمقارنة محتواه بمجال المحتوي الذي صمم لتقديره . وهذا المنهج الوصفي لا يزال ملائماً في أثبات صلاحية الاختبارات لتطبيقات معينة .

وعندما انتقلت عملية الاختبار إلي مرحلتها الثانية بمرور الوقت ، تحول الاهتمام إلي التنبؤ ، وكيف سيستجيب مختلف الأفراد في موقف معين الآن ، وفي زمن لاحق كيف يتفاعل الأفراد في مواقف مختلفة معينة ، وأصبح الأداء في الموقف المراد التنبؤ بالسلوك فيه يصمم كمحك. لهذا الغرض ، كان صدق الاختبار يذكر عادة كمعامل ارتباط بين نتائج الاختبار ومقياس مباشر ومستقل لهذا المحك . وهذا الأجراء ملائم بشكل خاص لاستخدام الاختبارات في اختيار أو وضع الأفراد فقي البرامج التعليمية ، وللوظائف ، أو لبرامج علاجية معينة . وهكذا ، فبالنسبة لاختبار القدرة الميكانيكية ، قد يكون المحك هو أداء العمل في وقت لاحق كميكانيكي ، ولاختبار القدرة الدراسية ، قد يكون المحك هو الصفوف الدراسية ، ولاختبار الأعصاب ، يمكن أن يكون المحك هو هوة ترتيب التداعيات أو المعلومات الاخري المتعلقة بسلوك الفرد في مواقف الحياة المختلفة .

وتعكس المرحلة الحالية في تاريخ الاختبار اتجاهين أساسيين هما :

أ- توجهاً نظرياً قوياً .

ب- ارتباط وثيق بين النظرية النفسية والتأكد من الصحة عن طريق الاختبار التجريبي للفرض .

وهذان الاتجاهان واضحان في إعداد الاختبار والتثبت من صحته ، كما هما واضحان في مجالات أخري لعلم النفس . وثمة نتيجة لهذين الاتجاهين تتمثل في زيادة الاعتراف بقيمة التركيب العقلي في وصف وفهم السلوك الإنساني ، والتركيبات العقلية هي فئات واسعة مستمدة من الجوانب المشتركة التي تشترك فيها المتغيرات السلوكية الملحوظة مباشرة . وأدي الاهتمام بالتركيب العقلي إلي إدخال ما كان يعتبر في البداية نوعاً ثالثاً من صدق الاختبار ، وبالتحديد صدق المنشأ أو التركيب . وأصبح صدق المنشأ أو التركيب فيما بعد يعرف بمفهوم " الصدق الأساسي الشامل لمقدار ما يحدد ما يقيسه الاختبار ".

ويعتبر المحتوي وإجراءات الصدق التنبؤي من بين المصادر الكثيرة للمعلومات التي تسهم في تعريف وفهم التركيبات العقلية التي يقدرها الاختبار . وفي نفس الوقت ، فأنها توفر معلومات قيمة في حد ذاتها ومهمة في تقييم الاختبارات لاستخدامات معينة . ومن ثم فإن المفاهيم ظلت باقية ، رغم دمجها في المفهوم الشامل لصدق المنشأ أو التركيب .

إجراءات وصف المحتوي :

سوف نتناول في إجراءات وصف المحتوي : طبيعتها ثم الإجراءات الخاصة والإجراءات التجريبية ثم الاستخدامات .

تضم إجراءات اثبات صحة وصف المحتوي الفحص النظامي لمحتوي الاختبار لتحديد ما إذا كان يغطي عينة ممثلة مجال السلوك المراد قياسه ، ويستخدم إجراء أثبات الصحة هذا بشكل عام في الاختبارات التي تصمم لقياس

درجة إتقان الفرد لمهارة معينة أو مقرر دراسي . وقد يبدو أن مجرد الفحص لمحتوي الاختبار يكون كافياً لإثبات صدقه بالنسبة لهذا الغرض . فاختبار الضرب ، الهجاء ، وإمساك الدفاتر علي التوالي . وليس الحل سهلاً كما يبدوا . فثمة صعوبة تتمثل في إعداد عينات كافية لمجال المفردة .

ويجب تحليل مجال السلوك المراد اختباره بشكل منتظم للتأكد من أن كل الجوانب الأساسية تغطيها مفردات الاختبار ، وبالنسبة الصحيحة . وعلي سبيل المثال ، فإن الاختبار يمكن أن يصبح محملاً بجوانب الميدان التي تساعد في أعداد المفردات الموضوعية ، ويجب وصف المجال محل الاعتبار بشكل كامل مقدماً ، بدلاً من تعريفه بعد إعداد الاختبار .

والاختبار التربوي جيد الإعداد ، علي سبيل المثل ، يجب أن يغطي أهداف التعليم ، وليس مجرد موضوعه . وعلي ذلك يجب تعريف المحتوي بشكل واسع لكي يشمل الأهداف الأساسية ، مثل تطبيق المبادئ وتفسير البيانات ، بالإضافة إلي معرفة الحقائق . وعلاوة علي ذلك ، فإن الصدق يعتمد علي ملائمة

استجابات الفرد لمجال السلوك محل الدراسة ، وليس علي الملائمة الواضحة لمحتوي المفردة . ومجرد الفحص للاختبار يمكن أن يفشل في الكشف عن العمليات المستخدمة بالفعل من قبل المفحوصين في اجتياز الاختبار .

ومن المهم أيضا الحذر من اي ميل للمبالغة في التعميم فيما يتعلق بالمجال الذي يقيسه الاختبار . علي سبيل المثال ، فان اختبار الاختيار من متعدد في الهجاء يمكن أن يقيس القدرة علي التعرف علي الكلمات التي تم تهجيها بشكل صحيح او بالخطأ . ولكن لا يمكن افتراض أن هذا الاختبار يقيس أيضاً القدرة علي التهجي الصحيح من الإملاء ، تكرار أخطاء الهجاء في التعبير التحريري ، وجوانب أخري من القدرة الهجائية . وهناك صعوبة أخري تظهر من إضافة عوامل

غير ملائمة في نتائج الاختبار . علي سبيل المثال ، فإن الاختبار المصمم لقياس الكفاءة في الرياضيات أو الميكانيكا يمكن أن يتأثر بلا داع بالقدرة علي فهم التوجيهات اللفظية أو بسرعة أداء المهام الروتينية البسيطة .

إجراءات خاصة :

يتم تحقيق صدق محتوي الاختبار بداية من خلال اختيار المفردات الملائمة . وبالنسبة للاختبارات التربوية ، يسبق إعداد المفردات دراسة دقيقة ومنظمة للمقررات والكتب المقررة ، وكذلك التشاور مع خبراء المواد الدراسية . وعلي أساس المعلومات التي تم جمعها بهذا الشكل ، يتم تحديد ووضع مواصفات الاختبار بالنسبة لكتابة المفردات . ويجب أن تبين تلك المواصفات مجالات المحتوي أو الموضوعات التي يجب تغطيتها ، والأهداف التعليمية أو العمليات التي يراد اختبارها ، والأهمية النسبية للموضوعات الفردية والعمليات .ويجب أن تبين المواصفات النهائية عدد المفردات من كل نوع والتي يجب إعدادها لكل موضوع . علي سبيل المثال ، فإن تقييم القدرة علي القراءة يمكن أن تشمل فهم المفردات في السياق ، الفهم الحرفي للمحتوي ، واستنباط استدلالات صحيحة من المعلومات المقدمة . ويمكن أيضاً أن يعزز المادة من مصادر مختلفة مثل المقالات ، القصائد الشعرية ، والمقالات الصحفية . ويمكن أن يغطي اختبارات الرياضيات مهارات الحساب ، حل المسائل التي تقدم شفاهه ، وتطبيق العمليات العلمية علي البيئات الجديدة وغير المألوفة . ويجب أن تضم مناقشة إثبات صحة المحتوي في كتيب اختبار التحصيل التعليمي وصفاً للإجراءات المتبعة في التأكد من أن محتوي الاختبار ممثل وملائم . وإذا شارك خبراء المناهج في عملية إعداد الاختبار ، فيجب تحديد عددهم ومؤهلاتهم المهنية ، وإذا ما تم استخدامهم كمحكمين في تصنيف المفردات ، يتعين ذكر التعليمات التي وجهت لهم ، بالإضافة إلي مدي الاتفاق بين المحكمين . ونظراً لأن محتوي المناهج والمقررات يتغير بمرور الوقت ،

فمن المرغوب علي وجه الخصوص إعطاء التواريخ التي تم فيها استثارة خبراء المواد ، ويجب تقديم المعلومات بالمثل عن عدد وطبيعة قدرات المناهج واستعرض الكتب المقررة ، بما في ذلك تاريخ النشر .

وهناك عدد من الإجراءات التجريبية :

يتم أتباعها بشكل عام لتكملة إثبات صحة المحتوي لاختبار تحصيلي . ويمكن مراجعة كل من النتائج الكلية والأداء في المفردات لتحديد التقدم . وبشكل عام ، يتم الإبقاء علي المفردات التي تبين أكبر زيادة في النسب المئوية للأفراد الذين ينجحون فيها من صف أدني إلي صف أعلي . وهناك إجراءات تكميلية أخري يمكن استخدامها إذا لزم الأمر ، وتشمل تحليل أنواع الأخطاء الشائعة في الاختبار وملاحظة طرق العمل المستخدمة بواسطة من يدخلون الاختبار . ويمكن أن تتم الخطوة الأخيرة باختيار الطلاب بشكل فردي مع إعطاء تعليمات لهم للتفكير بصوت عال عند حل كل مسألة . ويمكن مراجعة مساهمة السرعة بملاحظة عدد الأشخاص الذين يفشلون في إنهاء الاختبار ، ولاكتشاف التأثير المحتمل غير الملائم للقدرة علي قراءة التعليمات علي أداء الاختبار بالنتائج في اختبار قراءة وفهم النصوص ، فإن إعطاء الأسئلة بدون قراءة القطعة التي بنيت الأسئلة عليها سوف يبين عدد الأسئلة التي يمكن الإجابة عليها ببساطة من المعلومات السابقة لمؤدي الاختبار أو الإشارات غير الملائمة الاخري . عندما تدعم إجراءات أثبات صحة المحتوي بتلك الضوابط التجريبية مثل التي أوضحناها ، فإنها توفر أسلوباً كافياً لتقييم اختبارات التحصيل . وهي تتيح لنا الإجابة علي سؤالين يعتبران علي جانب من الأهمية بالنسبة لصدق اختبارات التحصيل التعليمي والمهني وهما :

1) هل يغطي الاختبار عينة مماثلة للمعرفة والمهارات المحددة ؟

2) هل اداء الاختبار متحرر بشكل معقول من تأثير المتغيرات غير الملائمة ؟

إن اثبات صحة المحتوي يلاءم بشكل خاص الاختبارات المرتبطة بالمجال . ونظراً لأن الأداء في تلك الاختبارات يفسر في ضوء معني المحتوي ، فمن الواضح أن ثبات صحة المحتوي يعتبر مطلباً أساسياً لاستخدامها الفعال . ورغم ذلك ، فإن انواعا أخري كدليل للصدق تكون ملائمة أيضاً للتقييم الكامل لفاعلية تلك الاختبارات .ويتم تطبيق إثبات صلاحية المحتوي أيضاً علي اختبارات مهنية معينة تصمم لاختيار الموظفين وتصنيفهم ، وهذا النوع دليل إثبات الصحة ملائم يكون الاختبار عينة عمل فعلية أو يتطلب نفس المهارات والمعرفة المطلوبة في العمل . في تلك الحالات يتعين تنفيذ تحليل دقيق للعمل لإيضاح التشابه الوثيق بين أنشطة العمل والاختبار .

وبالنسبة لاختبارات القدرة والشخصية ، يعتبر إثبات الصلاحية غير ملائم ، ويمكن أن يكون مضلاً في واقع الأمر . رغم أن اعتبارات الملائمة والتمثيل بالنسبة للمحتوي يجب أن تدخل في المراحل الأولي لإنشاء أي اختبار ، فإن إثبات الصلاحية اللاحق لاختبارات القدرة أو الشخصية يتطلب تأكيد تدريبي . أن هذه الاختبارات تحمل تشابهتً داخلياً اقل بمجال السلوك الذي نحاول فرزه من اختبارات التحصيل . نتيجة لذلك ، فإن محتوي اختبارات القدرة و الشخصية لا تفعل أكثر من مجرد كشف الفروض التي دفعت مصمم الاختبار لاختيار نوع معين من المحتوي لقياس سمة معينة . وهذه الفروض تحتاج للتأكيد التجريبي لإثبات صدق الاختبار . وإذا كانت اختبارات التحصيل تبني علي مقر دراسي ، فان اختبارات القدرة والشخصية لا تبني علي مقرر تعليمي معين أو مجموعة متشابهة من الخبرات السابقة يمكن أن يستمد منها محتوي الاختبار . ومن ثم ،ففي الاختبارات الأخيرة ، يحتمل أن يختلف الأفراد في طرق العمل أو العمليات السيكولوجية المستخدمة في الإجابة عن نفس مفردات الاختبار . وفي تلك

الظروف ، قد يكون من المستحيل تقريباً تحديد الوظائف السيكولوجية التي يقيسها الاختبار من فحص محتواه . علي سبيل المثال ، فإن خريجي المدارس يمكنهم حل مسألة علي أسس لفظية أو تجريبية ، بينما يمكن أن يصل الميكانيكي إلي نفس الحل علي أساس التصور الفراغي . وان الاختبار الذي يقيس التفكير الرياضي لدي طلبة المدرسة الثانوية يمكن أن يقيس الفروق الفردية فقط في سرعة الحساب عندما يعطي لطلبة الكليات .

الصدق الظاهري :

يجب عدم الخلط بين صدق المحتوي والصدق الظاهري ، فالصدق الظاهري ليس صدقاً بالمعني الفني ، فهو يشير لما يقيسه الاختبار بالفعل ، بل لما يبدو سطحياً في القياس . ويتعلق الصدق الظاهري بما إذا كان الاختبار يبدو صادقاً بالنسبة للمفحوصين الذين يتناولونه ، الأفراد الإداريون الذين يقررون استخدامه ، ويعتبر الصدق الظاهري نفسه جانباً مرغوباً من جوانب الاختبارات . علي سبيل المثال ، عندما يتم توسيع الاختبارات المصممة أصلاً للأطفال والتي يتم تطويرها لاستخدامها مع الكبار . فأنها كثيراً ما تتعرض للنقد بسبب عدم صدقها الظاهري . ومن المؤكد أنه إذا كان صدق الاختبار يبدو غير ملائم ، فإن النتيجة ستكون ضعيفة ، بصرف النظر عن الصدق الفعلي للاختبار . وفي اختبار الكبار علي وجه الخصوص ، لا يكفي أن يكون الاختبار صادقاً من الناحية الموضوعية . ويحتاج أن يتحقق الصدق بنجاح في المواقف العملية . ويؤثر الصدق الظاهري علي قبول الاختبارات في القرارات التشريعية والسياسية ، بالإضافة إلي تقييمه من جانب الجمهور أو الرأي العام .

ويمكن تحسين الصدق الظاهري بإعادة صياغة مفردات الاختبار علي أسس تبدو ملائمة ومقبولة في الموقف الخاص الذي سيستخدم . علي سبيل المثال ، إذا

تنم اعداد اختبار للتفكير الرياضي البسيط لكي يستخدم مع الميكانيكيين ، فيجب صياغة المفردات في ضوء عمليات الآلة ، وليس علي أساس مسائل الكتب المدرسية التقليدية . وبالمثل فإن اختبار الرياضيات لأفراد الأسطول يمكن صياغة مفرداته تمشياً مع المصطلحات البحرية ، بدون تغيير الوظائف المقاسة .

ومن المؤكد أن الصدق الظاهري لا يجب أن يعتبر كبديل للصدق الذي يتم تحديده بشكل موضوعي . ولا يمكن افتراض أن تحسين الصدق الظاهري للاختبار سوف يحسن صدقه الموضوعي . ويجب مراجعة صدق الاختبار في شكله النهائي بشكل مباشر .

إجراءات التبوء بالمحك :

إثبات الصحة التنبؤية المتزامنة :

تبين إجراءات إثبات الصحة التنبؤية المتزامنة مدي نجاح الاختبار في التنبؤ بأداء الفرد في أنشطة محددة . ويمكن الحصول علي المحك الذي علي أساسه تثبت صحة النتائج في نفس الوقت تقريباً الذي يتم فيه الحصول علي نتائج الاختبار أو بعد فترة محددة . وقد ميزت محكات الاختبار اثبات الصحة التنبؤية المتزامنة علي أساس العلاقات الزمنية بين المحك والاختبار ، ويمكن استخدام لفظ التنبؤ بالمعني الأوسع ، لكي يشير للتنبؤ بالمحك من الاختبار بأي موقف ، أو بمعني المحدود للتنبؤ في فترة زمنية معينة ، ويستخدم هذا اللفظ بالمعني الأخير في تعبير الصدق التنبؤي .

والمعلومات التي يوفرها إثبات الصدق التنبؤي تعتبر مهمة للاختبارات المستخدمة في وتصنيف الأفراد ، ويمثل توظيف طالبي العمل ، اختبار الطلاب للقبول بالكليات أو المدارس الفنية ، وتوزيع العسكريين علي برامج التدريب المهني أمثلة لنوع القرارات التي تتطلب معرفة بالصدق التنبؤي لاختبارات في فرز

طالبي الوظائف المحتمل أن تظهر عندهم اضطرابات عاطفية في البيئات الضاغطة ، واستخدام الاختبارات لتحديد المريض نفسياً الأكثر احتمالاً للاستفادة من علاج نفسي خاص .

وفي بعض الحالات ، يتم استخدام إثبات الصحة المتزامنة كمجرد بديل لإثبات الصحة التنبؤية .

وفيما يتعلق بالاستخدامات الخاصة للاختبارات السيكولوجية ، فأن التنبؤ المتزامن هو النوع الأكثر ملائمة ، ويعتمد الفرق المنطقي بين إثبات الصحة التنبؤية والصحة المتزامنة ليس علي زمن الاختبار بل علي أهداف الاختبار . واثبات الصحة التزامنية ملائم للاختبارات المستخدمة في تشخيص الوضع القائم ، وليس بالنتائج المستقبلية . ويمكن إيضاح الفرق بتوجيه السؤال (هل يصلح أحمد كطيار مقبول ؟) ، أو (هل يتمتع احمد بالمقدمات التي تجعل منه طياراً مقبولاً ؟) ، إن السؤال الأول يتطلب إثبات صحة تزامنية ، كما يتطلب السؤال الثاني أثبات صحة تنبؤية .

تشويش المحك :

ثمة تحذير أساسي في إيجاد صدق الاختبار يتمثل في التأكد من نتائج الاختبار في حد ذاتها لا تؤثر علي وضع محك . علي سبيل المثال ، إذا كان معلم المدرسة أو المشرف في مصنع يعرف أن فرد ما أحرز نتائج ضعيفة في اختبار قدرة ، فإن تلك المعرفة يمكن أن تؤثر علي التقدير الذي يمنح للطالب أو الترتيب الذي يخصص للعامل . وهذه الآثار يمكن أن ترفع الارتباط بين نتائج الاختبار والمحك بطريقة مصطنعة تماماً . وهذا المصدر المحتمل للخطأ في إثبات صحة الاختبار يعرف بتشويش المحك . حيث أن ترتيبات المحك تصبح مشوشة بمعرفة المقدر لنتائج الاختبار ، ولمنع ظهور هذا الخطأ ، من الضروري الا يعرف اي شخص من

الذين يشاركون في تحديد تقديرات المحك أي شئ عن نتائج اختبار الممتحنين . لهذا السبب ، فإن نتائج الاختبارات المستخدمة في فحص الاختبار يجب أن تظل سرية تماماً ، ومن الصعب أحياناً اقناع المدرسين ، أصحاب الأعمال ، الضباط العسكريين ، وغيرهم بأن ذلك الحذر ضروري ، وفي استعمالهم لكل المعلومات المتاحة عن القرارات العملية ، وهؤلاء الأشخاص يمكن أن يفشلوا في إدراك أن نتائج الاختبار يجب أن تنحي جانباً حتي تتضح بيانات المحك ويمكن مراجعة الصدق .

المحكات :

يمكن أثبات صدق الاختبار في مواجهة عدد من المحكات مساوي للاستخدامات الخاصة للاختبار . وأي وسيلة لتقدير السلوك في أي موقف يمكن أن توفر مقياس محك لبعض الأغراض الخاصة .

والمعايير المستخدمة في إيجاد الصلاحيات المذكورة في كتيب الاختبار . ومن بين المحكات الأكثر استخداماً في إثبات صدق اختبارات الذكاء هو مؤشر التحصيل الأكاديمي . لهذا اسبب وصفت تلك الاختبارات بدقة أكبر كمقاييس للقدرة الدراسية . وتشمل المؤشرات الخاصة المستخدمة كمقاييس محك لذكاء التقديرات الدراسية ، نتائج اختبارات التحصيل ، تقديرات المعلمين . وبمقدار ما تصبغ تلك التقديرات المعطاة في موقف أكاديمي بدرجة كبيرة بالأداء الدراسي للفرد ، يمكن تصنيفها بشكل صحيح بمحك التحصيل الأكاديمي ، وقد وفرت المؤشرات المختلفة للتحصيل الأكاديمي بيانات معيارية علي كل المستويات التعليمية ، من الصفوف الابتدائية حتي المدارس المتوسطة والمدارس الشاملة . ورغم استخدامها في الأساس في إثبات صدق اختبارات الذكاء العام ، فإنها استخدمت أيضاً كمحكات في اختبارات القدرات المتعددة

واختبارات الشخصية . وعند أثبات صدق أي نوع من هذه الاختبارات للاستعمال في اختيار طلبة لكليات ، علي سبيل المثال، فإن محكاً شائعاً هو متوسط تقديرات الطالب المستجد . وهذا المقياس يمثل متوسط التقدير في كل المقررات التي تمت دراستها خلال عام الطالب المستجد .

وثمة متغير لمحك التحصيل الأكاديمي كثيراً ما يستخدم مع الكبار خارج المدرسة يتمثل في مقدار التعليم الذي اتمه الفرد . وينتظر أن يواصل الأفراد الأكثر ذكاءً بشكل عام تعليمهم لفترة أطول ، بينما يتسرب الأفراد الأقل ذكاءً من المدرسة في وقت مبكر .

ويتمثل الفرق الكامن وراء هذا المحك في أن السلم التعليمي يعمل بشكل مستمر كتأثير انتقائي يلغي غير القادرين علي الاستمرار فيما وراء كل خطوة .

ولأغراض كثيرة ، فإن أفضل أنواع مقاييس المحك هي التي تبني علي سجلات المتابعة الخاصة بأداء العمل الفعلي . وتم استخدام هذا المحك بدرجة ما في إثبات صلاحية اختبارات الذكاء العام بالإضافة إلي اختيارات الشخصية ، وبدرجة أكبر في أثبات صدق اختبارات القدرات الخاصة ، وهو محك شائع الاستعمال في إثبات صلاحية الاختبارات الموصي عليها لأعمال خاصة . وقد يختلف العمل المقصود اختلافاً كبيرا في كل من المستوي والنوع . ومعم مقاييس أداء العمل رغم أنها لا تمثل معايير نهائية ، علي الأقل توفر معايير وسيطة جيدة لأغراض اختباريه كثيرة . وفي هذا الصدد ، فإنها تفضل علي سجلات التدريب . وعلي الجانب الأخر فان قياس أداء العمل لا يتيح ظروفاً متشابهة كما هو محتمل أثناء التدريب . علاوة علي ذلك ، ونظراً لأنه ينطوي في العادة علي متابعة أطول ، فان معيار أداء العمل يحتمل ان يشمل نقصاً ي عدد المشاركين المتاحين . وبسبب التباين في طبية الأعمال المماثلة اسمياً في مؤسسات مختلفة ، فإن كتيبات الاختبار التي تتناول بيانات الصدق مقابل

معايير العمل يجب ان تصنف ليس فقط مقاييس المحك الخاصة المستخدمة بل أيضاً مهام العمل الذي يؤديه الأفراد .

ويشمل أثبات الصلاحية بطريقة المجموعات المتناقضة بشكل عام محكاً إدارياً يعكس التأثيرات التراكمية الانتقائية غير المحكومة للحياة اليومية . ويبني ذلك المحك بشكل مطلق علي البقاء داخل مجموعة خاصة ووضع الفرد في تلك المجموعة ، علي سبيل المثال ، فإن صدق اختبار القدرة الموسيقية يمكن تحقيقه من خلال مقارنة النتائج التي يتم الحصول عليها من جانب الطلبة المقيدين في مدرسة موسيقي مع نتائج طلبة ليسوا من المدرسة العليا والمتوسطة ، وبطبيعة الحال ، فإن المقارنة بين المجموعات تكون علي أساس أي محك ، مثل التقديرات المدرسية ، الترتيبات ، أو أداء العمل علي أية حال تعتبر الجماعات المتميزة أصبحت متمايزة بالتدريج . من خلال تشغيل المطالب المتعددة

وتستخدم طريقة مقارنة المجموعات علي نطاق واسع في أثبات صدق اختبارات الشخصية . وعلي ذلك ، فعند إثبات صدق اختبار للسمات الاجتماعية ، فإن اختبار أداء أفراد المبيعات والمديرين التنفيذيين ممكن أن يقارن باختبار أداء الكتبة أو المهندسين . ويتمثل الغرض الكامن وراء هذا الأجراء في أنه بالإشارة لكثير من السمات الاجتماعية . فأن الأفراد الذين دخلوا وظلوا في تلك الوظائف كبائعين أو كمديرين سوف يتميزون كجماعة في تلك المجالات على الأفراد الذين في العمل الكتابي أو الهندسي . وبالمثل فان طلبة المدارس المتوسطة الذين اشتركوا في أنشطة كثيرة خارج المنهج يمكن أن يقارنوا بالطلاب الذين لم يشاركوا في أي أنشطة . وتم استخدام الجماعات المهنية في أعداد واثبات صدق اختبارات الكفاءة ، مثل اختبار الكفاءة المهنية ، وكذلك في أعداد مقاييس الاتجاهات ، وفي بعض الأحيان تستخدم جماعات اخري مثل الجماعات السياسية ،

الدينية ، الجغرافية أو غيرها من الجماعات الخاصة في إثبات صدق مقاييس الاتجاهات ، لتوفير أراء متنوعة عن مختلف القضايا .

وعند إثبات صدق الاختبارات مرجعية المجال تجريبياً ، يتم استخدام طريقة المقارنة بين المجموعات ، بالإضافة إلي الإجراءات المعتادة لإثبات صدق المحتوي ، لهذا الغرض فإن الجماعات التي تختلف في مقدار التعليم تقارن في أداء الاختبار ، وإذا استخدم تقدير الإتقان ، يمكن أجراء تحليل (2×2) ، وتتم فيه مقارنة نسبة نتائج الناجحين والراسبين في مجموعة ما بعد التعليم . ويمكن عمل مقارنات مماثلة إذا تم تطبيق الاختبار علي أطفال المدارس في صف أدني وصف أعلي دراسياً .

وعند أعداد اختبارات معينة للشخصية ، يستخدم التشخيص العلاجي كأساس لاختيار المفردات ، وكدليل علي صدق الاختبار في نفس الوقت ، ويمكن استخدام التشخيص العلاجي كمحك مقبول بشرط أن يبني علي ملاحظة مطولة وتاريخ حالة مفصل ، وليس علي مقابلة علاجية سريعة أو فحص .

ولقد تم استخدام التقديرات في أثبات صدق معظم الاختبارات ،، وهي مفيدة علي وجه الخصوص في توفير محكات لاختبارات الشخصية ، حيث إن المحكات الموضوعية يصعب الحصول عليها في هذا المجال . وينطبق ذلك بشكل خاص علي السمات الاجتماعية ، التي يمكن أن تشكل فيها التصنيفات المبينة علي الاتصال الشخصي المحك الأكثر قبولاً من الناحية المنطقية . ويمكن أن تخضع التقديرات للكثير من الأخطاء في الحكم ، ولكن عندما يتم الحصول ليها في ظروف محكومة بعناية فإنها تمثل مصدراً جيداً لبيانات المحك .

واخيراً ، فإن الارتباطات بين اختبار جديد والاختبارات المتاحة من قبل غالباً ما يستشهد بها كدليل علي الصدق ، وعندما يكون الاختبار الجديد شكلاً

مختصراً أو مبسطاً من اختبار متاح في الوقت الحالي ، يمكن اعتبار الاختبار الأخير في أحسن الأحوال كمقياس للمحك . وعلاوة علي ذلك ، فإن اختبار الورقة والقلم يمكن اثبات صدقه مقابل اختبار أداء مفصل ومطول . وقد استخدم اختبار ستانفورد . بينية علي سبيل المثال ، بشكل متكرر كمحك لإثبات صدق الاختبارات الجماعية .

وقد ركز التطور في بناء الاختبارات في الثمانينيات والتسعينيات علي الانتباه إلي تحليل المحك ، ويعتبر هذا الجانب في إعداد الاختبار لم يحظ بالاهتمام في بحوث الاختبارات التقليدية . ويمرور الوقت ، ارتفعت اصوات للتنبيه إلي الحاجة لبحوث في مجال المحكات ، وكن المضمون العملي كان ضئيلاً . وأصبح معروفاً علي نطاق واسع الآن . أن صدق الاختبار يمكن دراسته بفاعلية بتعريف التكوينات أو التركيبات الرئيسية في أداء العمل المعين ثم اختيار او إعداد اختبارات تقيم نتائجها التكوينات المطلوبة .

تعميم الصدق :

يستخدم صدق المحك . التنبؤ غالباً في الدراسات التي تتناول إثبات الصدق ، والتي يتم فيها تقييم فاعلية الاختبار بالنسبة لبرنامج خاص . وهذا هو المنهج المتبع ، علي سبيل المثال ، عندما ترغب كلية ما في تحديد مدي أمكانية اختيار القدرة الأكاديمية للتنبؤ بأداء الطلاب في مقرر دراسي . ويمكن وصف صدق المحك . التنبؤ بالصدق العملي لاختبار ما من اجل غرض محدد . وعند ربط اختيارات القدرة المقننة لأول مرة بالأداء في أعمال مماثلة في دراسات اثبات الصدق ، وجد أن معاملات الصدق تختلف اختلافاً كبيراً . ولوحظ تباين مماثل في المقررات الدراسية المختلفة . وهذه النتائج أدت إلي قلق واسع النطاق متعلق بقابلية تعميم صدق الاختبار علي المواقف المختلفة ، وحتى منتصف السبعينات ،

اعتبرت الخصوصية الموقفية للمقاومات السيكولوجية بشكل عام كقيد علي فائدة و استخدام الاختبارات المقننة في اختبار الأفراد . وفي تحليل إحصائي متطور للمشكلة تبين أن كثير من التباين بين معاملات الصدق التي تم الحصول عليها يمكن أن تكون ناتجة عن صغر حجم العينة ، عدم الثبات الخاص بالمحك ، وقيد النطاق في عينات الموظفين .

والعينات المتاحة لإثبات صدق الاختبار صغيرة الحجم لا تعطي تقديراً ثابتاً للارتباط بين عامل التنبؤ والمحك . ولنفس السبب ، فإن المعاملات التي تم الحصول عليها كانت منخفضة بحيث لا تصل إلي مستوي الدلالة الإحصائية في العينة المستخدمة ، وعلي ذلك يمكن أن تفشل في توفير الدليل علي صدق الاختبار .

وبتطبيق الأساليب المطورة علي البيانات التي تم الحصول عليها من عينات كبيرة مستمدة من التخصصات المهنية ، أمكن إثبات أن صدق اختبارات القدرة اللفظية ، والقدرة العددية ، والقدرة الاستدلالية يمكن أن تعمم علي المهن أكثر مما كان معروفاً من قبل .

ما وراء التحليل :

توفر الإجراءات الإحصائية في دراسة تعميم الصدق طريقة لدمج النتائج التي يتم الحصول عليها من دراسات مختلفة . ويمكن أن نجمع بيانات من البحوث القديمة والحديثة ، من دراسات أجريت في أماكن مختلفة أو من معلومات متاحة في دراسات منشورة .

وشهدت حقبة الثمانينيات والتسعينيات ثورة سريعة من بحوث ما بعد أو ما وراء التحليل في كل مجال من مجالات علم النفس ، ويحتمل أن يكون استخدامه في مجال الاختيار المهني وبحوث التصنيف قد جذب الانتباه ، ويزداد

الانتباه بما وراء التحليل بشكل مستمر ، ويتم تطوير تنقيحات إجرائية باستمرار . ورغم أن بعض الحديث يظل حول الأساليب الخاصة ، فأن النتائج الأساسية لا تختلف اختلافا كبيراً عبر الإجراءات .

طرق تحديد المنشأ أو البناء أو التكوين :

تم إدخال لفظ صدق المنشأ أو البناء رسمياً إلي معجم القياس النفسي عام (1954) وفي التوصيات الفنية للاختبارات النفسية والأساليب التشخيصية ، وظهر أول تفسير لصدق البناء في مقالة كرونباخ (1955) ، وركز إثبات صدق البناء علي دورة نظرية علم النفس في بناء الاختبار ، وعلي الحاجة لصياغة فروض يمكن إثباتها أو عدم إثباتها في ثنايا عملية إثبات الصدق . ويمثل البناء لاختبار المدى الذي يمكن أن يقلل عنده أن الاختبار يقيس البناء النظري أو السمة . وتتمثل أمثلة ذلك في القدرة الدراسية ، الفهم الميكانيكي ، الطلاقة اللفظة ، سرعة المشي ، العصاب والقلق . ويتم تطوير كل بناء لتفسير وتنظيم ثبات الإجابة موضع الملاحظة ، ويستمد من علاقات بينية راسخة بين المقاييس السلوكية . ويتطلب أثبات صدق المنشأ التراكم التدريجي للمعلومات من مصادر مختلفة وتمثل أي بيانات تلقي الضوء علي طبيعة السمة والظروف التي تؤثر علي تطور تلك السمة وإيضاحها دليلاً ملائماً علي إثبات الصدق . وسوف نتناول فيما يلي أساليب خاصة تسهم في تعريف البناء أو المنشأ .

التغيرات التطورية :

1) اختلاف عمر (التمييز أو المفاضلة علي أساس العمر الزمني):

ثمة محك رئيسي يستخدم في ثبات صدق العديد من اختبارات الذكاء التقليدية يتمثل في اختلاف العمر، ويتم ضبط تلك الاختبارات مثل اختبار ستانفورد . بينيه ومعظم اختبارات ما قبل المدرسة علي العمر الزمني لتقرير ما إذا

كانت النتائج تبين زيادة مضطردة مع تقدم العمر ، وحيث انه من المتوقع أن تزيد القدرات مع زيادة العمر خلال الطفولة ، فإن نتائج الاختبار يجب أن تبين تلك الزيادة بالمثل ، إذا كان الاختبار صادقاً. ومفهوم مقياس الذكاء هنا وكما استهله بينية يقوم علي فرض مؤداه أن الذكاء يزداد بزيادة العمر حتي مرحلة النضج علي الأقل .

أن محك اختلاف العمر، بطبيعة الحال ، لا ينطبق علي أي وظائف لا تظهر فيها تغيرات دقيقة وثابتة وفقاً للعمر الزمني . وعلي ذلك لا يصلح محك اختلاف العمر في مجال قياس الشخصية . ويري علماء القياس النفسي أن زيادة درجات الأفراد بزيادة أعمارهم لا تعطي دلالة أكيدة علي ما يقيسه الاختبار ، وعلي الجانب الأخر إثبات أن الاختبار يقيس شيئاً يزداد بزيادة العمر لا يحدد المجال الذي يغطيه الاختبار بدقة كبيرة .

وهناك نقطة أخيرة يجب تأكيدها وتتعلق بتفسير محك العمر . فالاختيار النفسي الذي يثبت صدقه بهذا المحك يقيس خواص السلوك التي تزيد بزيادة العمر في ظل الظروف السائدة في البيئة التي تم فيها تقنين الاختبار ، ونظراً لان الثقافات المختلفة يمكن أن تحفز وترسخ تطور سلوكيات مختلفة ، فلا يمكن افتراض أن محك اختلاف العمر محكاً شاملاً .

والتحليلات التطورية جوهرية في إثبات صدق البناء بالنسبة لمقاييس بياجيه في النمو العقلي . وثمة فرض أساسي لتلك المقاييس يتمثل في الترتيب التتابعي للتطور ، بحيث يكون تحقيق المراحل السابقة في تطور المفهوم طلباً أساسياً لاكتساب المهارات الخاصة بالمفاهيم اللاحقة . وهكذا هناك تسلسل داخلي في محتوي تلك المقاييس . وبناء علي ذلك ، فإن إثبات صدق البناء للمقاييس الترتيبية يجب أن يشمل بيانات تجريبية عن الثبات التتابعي للخطوات

المتتابعة . ويشمل ذلك ضبط اداء الأطفال علي مستويات مختلفة في تطور أي مفهوم يتم اختباره .

2) الارتباط باختبارات اخري :

في بعض الأحيان يتم الاستشهاد بالاختبار بين اختبار جديد واختبارات مماثلة سابقة كدليل علي أن الاختبار الجديد يقيس نفس مجال السلوك العام تقريباً مثل الاختبارات الاخري التي تسمي بنفس الاسم . ويستخدم الارتباط باختبارات اخري في إيضاح أن الاختبار الجديد متحرر نسبياً من أثر عوامل معينة غير ملائمة . علي سبيل المثال ، فان اختبار قدرة خاصة او اختبار الشخصية لا يجب أن يكون ارتباطه عال باختبارات الذكاء العام أو اختبارات القدرة الدراسية . وطبقاً لذلك فإن الارتباطات باختبارات الذكاء العام ، القراءة أو الفهم اللفظي تكون أحيانا دليل علي الصدق . وفي هذه الحالات فإن الارتباطات العالية يمكن أن تجعل الاختبار موضع شك ، والارتباطات المنخفضة قد لا تضمن أو تحقق الصدق .

3) التحليل العاملي :

أن التحليل العاملي الذي تم إعداده كوسيلة لتحديد السمات النفسية ، ملائم بشكل خاص لإجراءات أثبات صدق البناء أو التكوين . ويعتبر التحليل العاملي بمثابة أسلوب إحصائي منقح لتحليل العلاقات البينية لبيانات السلوك .

وقد تم تطوير أساليب إحصائية دقيقة لتحديد العوامل المشتركة . وفي عملية التحليل العاملي ، ينخفض عدد المتغيرات التي علي أساسها يمكن وصف اداء الفرد لعدد من الاختبارات الأصلية في عدد صغير نسبياً من العوامل أو السمات المشتركة . وثمة غرض رئيسي للتحليل العاملي وهو تبسيط وصف

السلوك بتقليل عدد متغيرات الاختبار إلي عدد قليل من العوامل أو السمات المشتركة .

ويعد تحليل العوامل ، يمكن استغلالها في وصف التركيب العاملي للاختبار . وعلي ذلك يمكن وصف كل اختبار في ضوء العوامل الرئيسية التي تحدد نتائجه ، مع وزن كل عامل وارتباطه بالاختبار . وأحياناً يذكر أن هذا الارتباط هو الصدق العاملي للاختبار . فإذا كان ارتباط عامل الفهم اللفظي باختبار المفردات اللغوية (0.66) ، فإن الصدق العاملي لاختبار المفردات اللغوية كمقياس لعامل الفهم اللفظي يكون (0.66) .

4) الاتساق الداخلي :

تتمثل الخاصية الأساسية لهذه الطريقة في أن المحك ما هو إلا الدرجة الكلية للاختبار نفسه . وفي بعض الأحيان يتم استخدام طريقة المجموعات المضادة علي أساس نتيجة الاختبار الكلية . وبعد ذلك يتم مقارنة المجموعة الأعلى في كل مفردة من مفردات الاختبار بأداء المجموعة الأدنى وتعتبر المفردات لا تظهر نسبة كبيرة من النجاح في المجموعة الأعلى وليس في المجموعة الأدنى غير صادقة ، وأما أن تحذف أو تعدل ويمكن استخدام الإجراءات الارتباطيه لهذا الغرض أيضاً . علي سبيل المثال ، فإن الارتباط الثنائي بين (ناجح - راسب) في كل مفردة ونتيجة الاختبار الكلية يمكن حسابها ، والمفردات التي تعطي ارتباطات دالة بين المفردة . الاختبار يمكن الاحتفاظ بها . وبذلك يمكن القول بأن الاختبار الذي تم اختيار مفرداته بهذه الطريقة سوف يظهر اتساقا داخلياً ، نظراً لان كل مفردة تفرق بين المجيبين في نفس الاتجاه كما يفعل الاختبار كله .

وهناك استخدام أخر لطريقة الاتساق الداخلي يتمثل في ارتباط نتائج الاختبار الفرعي بالدرجة الكلية . وكثيراً من اختبارات الذكاء . علي سبيل

المثـال ، تتكـون مـن اختبـارات فرعيـة تنفـذ بـشكل منفـصل (مثـل اختيـار المفـردات اللغويـة ، الرياضيات ، إكمـال الـصور ، وغيرهـا) . والـي يـتم تجميـع نتائجهـا فـي إيجـاد الدرجـة الكليـة للاختبـار ، وأي اختبـار فرعـي يكـون ارتباطـه بالدرجـة الكليـة منخفضاً جداً يـتم حذفـه . وتكـون ارتباطـات الاختبـارات الفرعيـة المتبقيـة كـدليل علي الاتساق الداخلي لـلأداة ككـل .ومـن الواضـح أن ارتباطـات الاتـساق الـداخلي ، سـواء بنيت علـي مفـردات او اختبـارات فرعيـة ، هـي بالـضرورة مقـاييس للتجـانس . ونظراً لأنها تساعد فـي وصـف مجـال الـسلوك أو الـسمة الـتي يقيـسها الاختبـار ، فـإن درجـة تجانس الاختبار لها بعض الملائمة بصدق البنـاء أو التكـوين .

5) الصدق التقاربي و المميز :

لكي يتم إيضاح صدق البناء أو التكوين ، يتعين علينا أن نبين ليس فقـط أن الاختبـار يـرتبط ارتباطـا عاليـاً بـالمتغيرات الاخـري الـتي يجـب أن يختلـف عنهـا ، ويطلق علي ارتباط الاختبار بمتغيرات تتسق معه (الصدق التقاربي) ، ويطلـق علـي عـدم ارتبـاط اختبـار الـتفكير الكمـي بنتـائج الطـلاب فـي مقـرر الرياضـيات مثـالاً للصدق التقـاربي ، وانخفـاض ارتبـاط اختبـار الـتفكير الكمـي باختبـار فهـم القـراءة مثالاً للصدق المميز .

والـصدق المميـز ملائـم بـشكل خـاص لإثبـات صـدق اختبـارات الشخـصية ، والتي فيها يمكن أن تؤثر متغيرات غير ملامة علي النتائج بطرق مختلفة .

رابعاً المعايير The Norms :

ناقشنا حتي الآن الشروط الأساسية للاختبار النفسي من حيث موضوعيته وثباته وصدقه ، وهـي الشروط الواجـب توافرهـا فـي أي اختبـار لكـي يصلح لقيـاس الصفة التي وضع من اجل قياسها .

ولكن ليس معني هذه الشروط ان الاختبار اصبح وسيلة تعطي درجة ، ذات معني للصفة المعينة ، فهو حقاً يعطي درجة ولكن ما معني هذه الدرجة ، ما معني مثلاً أن درجة " أحمد " في اختبار الحساب (70) أن وضع الدرجة بهذا الشكل لا يمثل شيئاً إلا إذا كان هناك المعيار الذي نقيس عليه هذه الدرجة ونحدد علي ضوئه إذا كانت مرتفعة أو منخفضة أو متوسطة تمثل المستوي العادي للأفراد اللذين هم في مثل سن " أحمد " وظروفه .

وهناك عدد من المعايير يستخدم لهذا الغرض ، ويتم عن طريقها مقارنة بين درجة الفرد بدرجات بأفراد آخرين طبق عليهم نفس الاختبار .

ومن هنا تأتي أهمية تحديد خصائص العينة التي استخدمت في استخراج معايير الاختبار ، حتي تكون مطابقة درجة الفرد علي هذه المعايير سليمة ، فإذا كانت العينة المستخدمة في استخراج معايير اختبار ما مأخوذة مثلاً من بين طلبة المدارس الثانوية ، فأنها تكون صالحة فقط لأفراد هذا المستوي التعليمي ولا تصلح لمقابلة درجة طفل في المدرسة الابتدائية مثلاً .

وفي الحالات التي يكون الاختبار فيها صالحا للاستخدام لأكثر من عمر أو لأكثر من مستوي تعليمي يجب انه تتدرج المعايير حسب مستويات العمر أو الدراسة أو غيرها ، فيكون لكل عمر أو مستوي معيار تقاس عليه درجات أفراد العمر الواحد أو المستوي الواحد .

وهناك ثلاثة أنواع أساسية من المعايير تستخدم لتفسير درجات الاختبار هي :

1) معيار العمر .

2) المئينيات .

3) الدرجات المعيارية . ونوضح فيما يلي هذه الأنواع الثلاثة .

1. معيار العمر :

يستخدم هذا النوع من المعايير مع أية صفة تنمو بزيادة العمر ويكون المعيار في هذه الحالة هو متوسط درجات عينة الأفراد من نفس السن . ولا كانت هذه القيمة المتوسطة (أو المعيار) تزداد بزيادة العمر ، فسيصبح لكل عمر معيار مختلف تزداد قيمته من سن إلي أخر .

فمعيار الطول لسن ست سنوات هو متوسط طول عينه ممثلة لأطفال هذا السن . ومعيار الطول لسن سبع سنوات ، وهو اكبر بالطبع من المعيار السابق (الخاص بسن ست سنوات) وهكذا .

فإذا أردنا أن نعرف ما إذا كان احد الأطفال قصيراً أو متوسط الطول أم طويلاً ما علينا إلا أن نقارن طوله بالمعيار الذي يمثل سنه السابق (الخاص بست سنوات) .... وهكذا فإذا أردنا أن نعرف ما إذا كان أحد الأفراد قصيراً أم متوسط الطول أم طويلاً إلا أن نقارن طوله بالمعيار الذي يمثل سنه .

وقد استخدم هذا النوع من المعايير في كثير من الاختبارات وخاصة اختبارات الذكاء وأشهر استخداماته بهذا الخصوص هو استخدامه في اختبار ستانفورد . بينيه والاختبارات الاخري المعدلة لهذا الاختبار الأصلي ، لتحديد العمر العقلي .

العمر العقلي Mental Age :

من المعروف ان الذكاء خاصية تنمو بتقدم الأفراد فى العمر حتي يصل إلي سن معين . وقد استفاد بينيه من هذه الخاصية في وضع اختباره وكان الأساس الذي اعتمد عليه هو أننا إذا عرضنا مستوي النمو العقلي للطفل العادي

(متوسط النمو العقلي للأطفال الذين هم في نفس السن) ، أمكننا أن نحدد بالنسبة إلي هذا المستوي مدي تقدم أو تأخر اي طفل أخر من ناحية قدرته العقلية .

ومن اجل هذه الغاية وضع بينية مجموعات من الأسئلة وحدد لكل عمر من الأعمال العقلية مجموعة معينة تمثل متوسط النمو العقلي في هذا العمر وتدل عليه .

ولتحيد درجة أي طفل في هذه الاختبار ، يعطي الاختبار ليجيب عن أسئلته مبتدأ بالأسئلة التي تمثل الأعمال الأولي حتي يصل إلي مستوي معين ولا يستطيع أن يتعداه هذا المستوي الذي يحدد العمر العقلي للطفل فإذا أجاب طفل عن أسئلة سن (7) مثلاً يقال أن عمره العقلي (7) . فإذا كان عمر الطفل الزمني (7) أيضاً كان هذا الطفل عادياً . أو متوسط من حيث الذكاء أما إذا كان عمره (8) أو (9) فأنه يكون متأخر في الذكاء ، وإذا كان عمره الزمني (6) أو (5) فإن ذكاؤه يكون مرتفعاً .

وللتخلص من المشكلة من العلاقة بين العمر الزمني والعمر العقلي اقترح " وليام شترن "W,Stern قسمة العمر العقلي علي العمر الزمني للحصول علي النسبة العقلية . فإذا ضربا الناتج في (100) فأننا نحصل علي نسبة تحدد ذكاء الفرد وتسمي هذه النسبة نسبة الذكاء اي أن :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

هذه النسبة هي التي تحدد مستوي ذكاء الفرد ، وتبعاً لها يكون مستوي ذكاء الفرد العادي او المتوسط هو (100) فإذا ارتفعت النسبة عن (100) ارتفع الذكاء ، وإذا انخفضت عن (100) كان الذكاء منخفضاً.

فإذا كان لدينا ثلاث أطفال عمرهم العقلي (6)العمر الزمني للأول (4) ، والثاني (6) ، والثالث (8) فإن نسبة ذكاءهم ستكون :

نسبة ذكاء الأول = $\frac{6}{4} \times 100 = 150$.

نسبة ذكاء الثاني = $\frac{6}{6} \times 100 = 100$

نسبة ذكاء الثالث = $\frac{6}{8} \times 100 = 75$

بيد أننا يجب أن نشير إلي أن النمو العقلي لا يساير النمو الزمني تماماً ، ومن هنا لا تبقي النسبة ثابتة كما هو مفروض أن تكون. فضلاً عن أن الذكاء لا يستمر في النمو طيلة حياة الإنسان ، بل أنه يتوقف عند سن معين (عند حوالي 17 سنة تقريباً) . فإذا طبقنا معادلة نسبة الذكاء بعد هذا السن ستلاحظ تدهوراً مستمراً في ذكاء الفرد ، ولذلك يحسب العمر الزمني للفرد بعد هذه السن بحساب أنه (17) سنه ، مهما تقدم الفرد في العمر . لن نمو ذكاءه قد يتوقف عند حد معين فيجب أن نفترض حدا ثابتا لنموه الزمني كذلك هي السن التي توقف عندها نمو الذكاء ، حتي تكون المعادلة أكثر تمثيلاً لذكاء ه الحقيقي .

ولهذه الأسباب وغيرها يفضل اغلب المشتغلين بالقياس العقلي استخدام المعايير المئينية.

2. المئينيات Perecentiles

يعتبر المئين من أكثر المعايير استخداماً وخاصة في اختبارات القدرات العقلية ، وهو يحدد النسبة المئوية للحالات التي تقع بعد درجة معينة . فإذا كانت درجة الفرد في الاختبار هي (60) درجة ، وحصل (80٪) من افراد عينة مماثلة له علي درجات أقل من (60) ، تكون درجته المئينية هي (80) وهذا معناه انه افضل من (80٪) من أفراد العينة المماثلة .

وطبيعي أنه يجب أن يكون هناك أكثر من معيار مئيني للاختبار الواحد إذا كانت الصفة التي يقيسها الاختبار تتأثر بالسن أو المستوي التعليمي أو نحو ذلك ، يمثل كل معيار منها أحدي سنوات العمر المتتالية التي يستخدم الاختبار في حدودها ، أو احد المستويات التعليمية ... الخ حتي تقاس درجة الفرد علي المعيار المئيني الذي يمثل الذي يمثل سنه أو مستواه أو نحو ذلك .

وللحصول علي المئين نستخدم نفس الطريقة المستخدمة في تعيين الوسيط و الارباعي فكما قسمنا المجموعة إلي قسمين لتحديد النقطة الوسطي في حالة الوسيط وكما قسمناها إلي أربعة أجزاء في حالة الارباعي ، يمكن أن نقسمها إلي مائة جزء في حالة المئين . ويكون المئين هو النقطة التي تحدد هذه الأجزاء فالمئين (10) هو النقطة التي يقل عنها(10٪) من القيم ، والمئين (20) هو النقطة التي يقل عنها (20٪) ، والمئين (50) هو النقطة التي يقل عنها (50٪) من القيم (هو الوسيط)..... الخ .

ومتى تحددت القيم المقابلة لكل مئين فإنها توضع عادة في شكل جدول للاستفادة منها في معرفة المئين التي يقابل أي قيمة يعطيها الاختبار وبهذه الكيفية نستطيع تحديد ترتيب الفرد بالضبط بالنسبة للمجموعة .

وسهولة استخدام هذا النوع من المعايير ، فضلاً عن تحديده ترتيب الفرد ضمن المجموعة بدرجة كبيرة من الدقة ، هما السبب في انتشار استخدام المئينيات في الاختبارات النفسية المختلفة التي يلائمها هذا النوع من المعايير .

3. الدرجة المعيارية :

يمكن أن تستخدم هذه الدرجة المعيارية أيضاً كمعيار لمقارنة درجة الفرد بالنسبة لدرجات المجموعة التي هو عضو فيها . فإذا كانت الدرجة التي حصل عليها تلميذ في مادة ما هي (60) من المائة ، فإن هذه الدرجة (كما سبق أن رأينا) لا تدل في حد ذاتها عما إذا كان التلميذ متأخراً أو متقدماً بالنسبة لتلاميذ الفصل ، وقد يساعد علي معرفة مركز هذا التلميذ نسبياً ، أن نعرف متوسط درجات تلاميذ الفصل . فإذا كان المتوسط (50) فأن درجة التلميذ تقع في النصف العلوي من الدرجات . ولكننا لا نخرج من معرفة المتوسط بأكثر من ذلك .

أما إذا أردنا أن نعرف مركز التلميذ بالضبط ضمن النصف العلوي الذي تقع فيه الدرجات عن المتوسط بمقياس من مقاييس التشتت كالانحراف المعياري . وبذلك نحصل علي الدرجة التي يحصل عليها الفرد في الاختبار والمتوسط أي الانحراف المعياري أو :

$$\text{الدرجة المعيارية} = \frac{\text{الدرجة} - \text{المتوسط}}{\text{الانحراف المعياري}}$$

والدرجة المعيارية قد تكون سالبة إذا كانت الدرجة الأصلية أقل من المتوسط ، أو موجبة إذا كانت أكبر من المتوسط ، أما المتوسط فدرجته المعيارية صفر .

وتتضح قيمة الدرجة المعيارية التي يحصل عليها الفرد إذا عرفنا أن التوزيع الاعتدالي للدرجات تكون القيم محصورة بين (+3ع ،- 3 ع) وفيه تقع :

68.26 من الحالات بين +1ع ، - 1 ع .

95.44 من الحالات بين +2ع ،- 2 ع .

99.72 من الحالات بين +3 ع ، - 3 ع .

فإذا كانت درجة التلميذ كما ذكرنا هي (60) وكان المتوسط (50) ، وكان الانحراف المعياري للدرجات هو 9.5 فإن :

$$\text{الدرجة المعيارية} = \frac{60 - 50}{9.5} = 1.05$$

أي انها اكثر من (1 ع) .

وهذا معناه أن التلميذ الحاصل علي هذه الدرجة يتفوق علي كل التلاميذ الذين يقلون عن هذا الحد ويمثلون حوالي 84٪ من التوزيع . أما إذا كانت درجة التلميذ (70) فان :

$$\text{درجته المعيارية تصبح} = \frac{70 - 50}{9.5} = 2.10$$

ومعني هذه الدرجة انه يتفوق علي 98٪ من تلاميذ الفصل وهكذا .

خامساً سهولة استخدام الاختبار :

المقصود بسهولة استخدام الاختبار ان تكون إجراءات التطبيق سهلة وعملية ، وان تكون تعليماته واضحة تماماً وغير معقدة ، وان تكون له طريقة محددة للتصحيح . ولتحقيق هذه الغايات يفضل :

1. ان يكون للاختبار دليل Manual يعطي تعريفاً للاختبار من موضوعه والأغراض التي يفيد فيها استخدامه ، ومحتوياته الأساسية ، ويعطي فكرة عن عينة الأفراد التي استعين بها في استخراج ثوابت الاختبارات كمعاملات الصدق والثبات .... الخ ، والتي يصلح استخدام الاختبار من ثم معها ويتضمن الدليل كذلك المعايير الخاصة بالدرجات .

2. أن يمهد للاختبار بتعليمات واضحة محددة توضح المقصود من كل جزء من أجزاءه وكيفية الإجابة عن اسئلته ويحسن أن تتضمن التعليمات نموذجاً او اكثر لهذه الأسئلة وطريقة الإجابة عنها (طريقة وضع العلامات او الكلمات ..... الخ) في الجزء المخصص لها في ورقة الإجابة ، والأفضل ان تكون للاختبار ورقة إجابة منفصلة ، لسهولة وضع العلامات أو نحوها ولسهولة عملية التصحيح أيضاً .

3. يجب أن ترتب الأسئلة بشكل يسمح بتتبعها بسهولة والإجابة عنها باطراد ومن غير تعثر كأن ترتب الأسئلة حيث درجة صعوبتها فتوضع الأسئلة في مجموعات تختص كل مجموعة منها بموضع معين أو بجزء معين إذا كان الاختبار يشتمل علي اكثر من جزء او اكثر من موضوع او في صورة مجموعات من الصيغ ، إذا كانت أسئلة الاختبار موضوعة في اكثر من صيغة ، اسئلة التكميل مع بعضها مثلاً وكذلك أسئلة الصواب والخطأ وأسئلة المقابلة .....: او غير ذلك .

4. ان يكون هنـاك نموذجاً للإجابـة الصحيحة (مفتـاح key) وان يكـون هناك النموذج أو المفتاح سهل الاستعمال عملياً حتي لا تستغرق عملية التصحيح وقتاً كبيراً أو جهدا ، وحتى يتيسر لأي شخص أن يقوم بعملية التصحيح .

ومن التسهيلات التي تمثل هذه الغاية أن يكون المفتاح مثقباً يمكن وضعه علي ورقة الإجابة ثم نعد العلامات الصحيحة التي تظهر من ثقوب المفتاح ، وان يتم التصحيح عن طريق الآلات ..... إلي غير ذلك من التسهيلات التي تفيد بصفة خاصة في الحالات التي يطبق فيها الاختبار علي نطاق واسع وعلي عينات من مئات أو الاف الأفراد .

5. التكلفة :

ولا ينبغي أن يهمل الباحث العنصر الاقتصادي عند اختياره للاختبارات التي سيستخدمها .

الفصل الرابع

" بعض عوامل الموقف الاختباري "

مقدمة :

يتأثر أداء الطلاب للاختبارات ببعض العوامل منها يرجع للمفحوص منها اعداد المفحوص لأخذ الاختبار، وقلق الاختبار والحكمة الاختبارية ، ودوافع المفحوص والحاجة الجسمية والقدرة العقلية والغش في الامتحانات وتدريب المفحوصين علي أخذ الاختبار وهناك العوامل المتعلقة بالاختبار مثل صعوبة المفردات وترتيب المفردات ومعرفة ترتيب المفردات ، وطول الاختبار وتعليمات الاختبار ، وزمن الاختبار ونوع الاختبار ، وطرق تقدير الدرجات والتخمين والتصويب من أثر التخمين ، واختلاف أوزان الفقرات وسوف نوضح فيما يلي هذه العوامل .

أولاً إعداد التلاميذ لأخذ الاختبار :

يجب الاهتمام بأعداد التلاميذ لأخذ الاختبار ، فيجب أن يعلم التلاميذ انهم سوف يختبرون في وقت لاحق وقبل موعد إجراء الاختبار بوقت كاف حتي يستعدوا وإذا كان الهدف من الاختبار إثارة دافعية التلاميذ للاستذكار ، وتوجيهه جهودهم نحو التعلم فلا يجب أن يعرفوا أنهم سوف يختبرون فحسب بل يجب أن يعرفوا الجزء من المقرر الذي سوف يختبرون فيه حتي يمكنهم الاستعداد بصورة محددة ويتطلب هذا من المعلم أن يقوم بتخطيط عملية الاختبار قبل أن يبدأ تدريس المقرر كجزء من التخطيط العام للتدريس المقرر .

ومَن المهم أن يرغب المفحوص في أن تكون درجة الاختبار صادقة وصحيحة ذلك لأن هذه الرغبة ضمان لأن تكون الدرجة صادقة صحيحة لأن هذه الرغبة

ضمان لأن تكون هذه الدرجة صادقة فعلاً ويساعد التعاون بين المختبر والمفحوص علي إيجاد هذه الرغبة ولهذا كان أمراً يستحق أن يحرص المختبر علي توفره وهو ليس عسير التحقيق علي أي حال ، علي أنه من المهم أن يكون واضحاً أ، التعاون الذي نقصده بين المختبر والمفحوص ليس معناه أن تحال المسئولية كلها وتنتقل إلي المفحوص وإنما نقصد أن يجعل المختبر المفحوص يحس باستمرار بأن مشاركته في العمل مشاركة العضو العامل في أي جماعة ، ويجب أن يصور المختبر للمفحوص الاختبار علي أنه فرصة طيبة وأداءه عملية ميسرة يعرف عن طريقها شيئاً عن ويجب أن تقدم الاختبارات بتعليمات واضحة وبسيطة ومن الأفضل اختيار مدي فهم العينة التي ستختبر لهذه التعليمات واختبار مدي فهم التعليمات وصياغة البنود تصبح من الأمور الهامة في الإعداد النهائي للاختبار .

هناك العديد من وجهات النظر التي تناولت تعريف القلق ومنها ما يلي :

يري " حامد زهران " (1978) أن القلق هو حالة توتر شامل ومستمر نتيجة توقع تهديد خطر فعلي أو رمزي قد يحدث ، ويصاحبها خوف غامض وأعراض نفس جسمية .

ويعرف " احمد عكاشة " (1980) القلق علي أنه شعور عام غير سار بالتوقع والخوف والتوتر مصحوب عادة ببعض الإحساسات الجسمية خاصة زيادة نشاط الجهاز العصبي اللاإرادي يأتي في نوبات تتكرر في نفس الفرد ، وذلك مثل الشعور بالفراغ في فم المعدة ، أو السحبة في الصدر او الضيق في التنفس أو الصراع ، أو كثرة الحركة .

أما " احمد عبد الخالق " (1987) فيعرفه بانه انفعال غير سار وشعور مقدر بتهديد أو وهم أو عدم راحة وهو كذلك إحساس بالتوتر والشر والخوف الدائم الذي لا مبرر له من الناحية الموضوعية .

اما " انتصار يونس " (1978) فتعرف القلق بأنه عدم الارتياح والاستقرار الذهني الفزع الغامض والتوتر الزائد ، وهو كثير الحدوث في حياتنا اليومية فالإنسان لا يخلو من القلق فهو يعيشه لبناء مستقبله ويضع لنفسه الأهداف ، ويصعب عليه حلها ينتابه القلق ولكنه قلق طبيعي بل أكثر من ذلك أن القلق في شكله البسيط يؤدي إلي الحرص والحيطة .

وتتمثل نوبة القلق في اضطراب القلب واضطراب النفس ، والعرق والفزع ودوران الرأس والدخان والإسهال ، وكثرة إفراز البول واختلال الرؤية والشعور العام بالتعب فالأعراض الجسمية لعصاب القلق قد تشمل أياً من أجهزة الجسم أو كلها .

أ- القلق كحالة والقلق كسمة :

1. القلق كحالة :

عرفها " كاتل " (1961) بأنها حالة انفعالية عارضة أو عابرة للكائن البشري تتسم بالمشاعر الموضوعية التي يعيها الفرد ويسودها التوتر والخوف من شئ مرتفع وارتفاع في نشاط الجهاز العصبي الذاتي (اللاإرادي) كما تختلف حالة القلق عبر الوقت وقد أشاد كاتل إلي أن حالة القلق تتغير حسب الموقف وأن التباين بين تلك المواقف أكثر ارتفاعا من التباين بين الأفراد .

2. القلق كسمة :

عرفها " سبلبرجر" القلق بأنه عبارة عن استعداد سلوكي مكتسب يظل كامناً حتي ينبه وينشط منبهات داخلية أو خارجية تثير حالة القلق وقد أوضح كل من . كاتل و سبلبرجر. (1974) ان سمة القلق تشير إلي الفروق الفردية

المستقرة الثابتة التي تميز الأفراد في الميل إلي الاستجابة في لمواقف التي يدركونها علي أنها مهددة .

ب- قلق الاختبار :

قلق الاختبار هو نوع من القلق المرتبط بمواقف الاختبار بحيث تثير هذه المواقف في الفرد شعور بالخوف والهم العظيم عند مواجهة الاختبارات .

ويعتبر موضوع قلق الاختبار من المجالات التي يجب الاهتمام بدراستها فغي مصر نظراً لأن هذا المجال قد نال قسطاً كبيراً من الاهتمام في البحوث الأجنبية .

ولا يقتصر قلقل الاختبار علي الأطفال بل يمتد إلي الراشدين ويشير " سيبز " في عرض للبحوث الخاصة بقلق الاختبار إلي احتمال أن يكون دافعية الاختبار بالنسبة للشخص احد العوامل الهامة في إثارة هذا القلق ، وهو ما يعني أن الاختبار بوصفه اختبارا ، ودون النظر إلي مضمونه ما يقيسه يحمل دلالات معينة للمفحوص فقد يعكس شخص ما هذا القلق باعتباره ميكانيزماً دفاعياً لجأ له الفرد ليتخلص به من الالتحاق بالوظيفة التي يخشى الفشل فيها ، فدلالة الاختبار بالنسبة لهذا الشخص مرتبطة بأنه المنفذ للوظيفة .

ج- تعريف قلق الاختبار :

لقد تعددت المفاهيم التي تناولت قلق الاختبار ومنها ما يلي :

يعرفه johner - Hills(1971) علي انه مستوي نادر واستثنائي من الإثارة مرتبطة بأخذ الاختبار في الموقف الاختباري .

ويعرف " سبلبرجر Spilberger (1980) قلق الامتحان علي انه سمة في الشخصية في موقف محدد ويتكون من الانزعاج Worry والانفعالية

Emotion ويحدد الانزعاج علي انه اهتمام معرفي بالخوف من الفشل ،وتحدد الانفعالية علي انها ردود أفعال للجهاز العصبي الاوتونومي .

أما " فتحي عبد الحميد " (1987) فيعرفه علي أنه حالة انفعالية تتميز بالخوف من نتيجة الاختبار وينشأ القلق في موقف الاختبار ويعد منبأ له .

ويعرف " إيفا نس وماكندليس " قلق الاختبار بأنه توتر انفعالي يرتبط بالتقدم الأكاديمي ويواجه ويمارس بصورة عامة من قبل الاختبار في مستهل حياتهم الدراسية ثم يتناقص عادة عبر سنوات الطفولة المتوسطة وما بعدها باستثناء أولئك الذين يكونون أبطأ بتعلمهم بالمقارنة بزملائهم .

ولقد قدم " طارق محمود " (1984) تعريفاً إجرائيا للقلق المدرسي في المرحلة الثانوية علي النحو التالي القلق المدرسي هو مظهر علني أو صريح من الهموم والحالات غير السارة المرتبة بالمواقف المدرسية ، تتضمنه العلاقات بين الطالب ودرسيه وردود الأفعال الناجمة عن المواقف التحصيلية متمثلة في الخوف والفشل من قلق الامتحان .

أهمية دراسة قلق الاختبار

ويوضح لنا " روبرت كابلن " (1982) أهمية قلق الاختبار بقوله نحن نعتقد أن قلق الاختبار سوف يظل مجالا فعالا للبحث لثلاثة أسباب هي :.

1. أن قلق الاختبار هو علاقة كبيرة بكثير من النظريات العامة في علم النفس .
2. أن قلق الاختبار يعتبر مشكلة حقيقية لكثير من الطلاب .
3. وجود كثير من الدراسات الخاصة بقلق الاختبار التي تساعدنا في الوصول إلي فهم القلق بصفة عامه .

فقلق الاختبار يعتبر مشكلة حقيقية تواجه كثير من الطلاب وهي مصدر قلق ليس للطلاب فقط ، ولكن للأسرة كلها نظراً لأن الامتحانات في مدارس مصر وجامعاتها تشوبها كثير من العيوب التي تجعلها أداة تقويم أو توجيه غير ملائمة .

هـ - أنواع قلق الاختبار :

حدد " مونز سيمونز " أنواع قلق الاختبار باستخدام قلق التحصيل فيما يلي :

1. قلق ميسر وهو الذين لديهم درجات مقياس القلق الميسر أعلي من درجات مقياس القلق المعوق أي أن القلق الميسر – القلق المعوق = فرقاً موجباً .
2. قلق معوق وهو الذين لديهم مقياس القلق الميسر – القلق المعوق =فرقاً سالباً وكانت النسبة المئوية لمجموعة القلق الميسر 25٪ من العينة الكلية ونسبة مجموعة القلق المعوق 25٪ من العينة الكلية أما 50٪ جمعت درجاتهم في مقياس قلق الاختبار الميسر والمعوق وصنف علي أساس الطلاب الذين درجاتهم فوق المتوسط علي أنهم ذووا تأثير فعال بينما الطلاب درجاتهم دون المتوسط عرفوا علي أنهم عديموا التأثير .

والنوع الأول من القلق يعمل علي حفز المفحوص وجعله في حالة يقظة وتنبه بينما يؤدي الثاني علي هذا القدر من الوهن في أداءه وقد وجد " ألبرت و هابر "، اللذان يميزان بين هذين النوعين من القلق ما يؤيد هذا التفسير فالمفحوصين الذين يؤكدون أن العصبية أثناء الإجابة علي الاختبار تساعد علي تقديم أداء أفضل ارتبطت درجاتهم إيجابياً بين هذا القياس للقلق الميسر وبين أدائهم بينما ارتبطت سلباً درجة القلق السلبي او المثبط وأن كانت النتائج التي خرج بها هذان الباحثان أو غيرهما غير حاسمة .

و- القلق الاختباري الجنسي :

لقد وجد أن قلق الاختبار له ارتباط سالب بالأداء في المدرسة أو الكلية ووجد أن الارتباط السالب يتزايد بصورة أكبر عند الطالبات عن عند الذكور وارتفاع درجة القلق لدي البنات تعني أن البنات أكثر تحملاً للقلق عن البنين والبنات أكثر توافقاً للقلق من البنين وأن هذا الفرق من الممكن أن يؤثر في طبيعة العلاقة بين القلق والأداء العقلي لكلا الجنسين .

ولقد أكد سبلبرجر (1980) علي أن الإناث كن أعلي من الذكور في مستوي القلق بنسب تتراوح بين (3 : 5) درجات علي مقياس قلق الامتحان " لسبلبرجر "وذلك علي عينات شملت طلبة كليات وطلبة المدارس الثانوية وكذلك وجد أن " أن متوسط الدرجات التي حصل عليها الذكور بالنسبة للدرجة الكلية لقلق الامتحان أكثر ارتفاعا من متوسط الدرجات التي حصل عليها الذكور من طلبة الكليات .

ز- قياس قلق الاختبار :

تختص مقاييس قلق الاختبار بأنواع معينة من المواقف يمكن استخدامها بصورة أوضح في هذه المواقف عن المقاييس العامة للقلق ، وقد أكدت دراسات " ساراسون "(56- 1958) التي أجريت علي طلاب الجامعة أن مقاييس القلق العام ليس لها علاقة بالأداء العقلي ويوجد اتجاهين في قياس قلق الاختبار علي أساس مقاييس :

أ- قلق ميسر .

ب- قلق معوق .

واتجاه يقيس قلق الاختبار ويعطي درجة كلية للقلق .

ويـشير " فـاروق عبـد الفتـاح ط(1990) إلـي نتـائج بعـض الدراسـات والملاحظات العامة التي أجريت علي الجنس البشري وهذه النتائج هي :

1) يوجد ارتباط سالب بين مستوي القدرة ومستوي قلق الاختبار ولذا يبدوا أن الأفراد ذوي القدرة المرتفعة اقل قلقاً عندما يواجهون الاختبار .

2) يوجد ارتباط موجب بين مستوي القلق ومستوي الطموح فهؤلاء الذين يبدوا أكثر طموحاً عندما يواجهون الاختبار يبدون أكثر رغبة أو أكثر حاجة للإجابة بشكل طيب .

3) الدرجات المتوسطة من القلق تؤدي إلي تحسن الأداء في الاختبار ولكن الدرجات المتطرفة يبدو أنها تعوق الأداء في الاختبار .

4) كلما اعتاد التلميذ علي نوع معين من الاختبارات فإنه يكون اقل احتمالا للوقوع فريسة للقلق الشديد عندما يواجه هذا النوع من الاختبارات .

5) يمكن أن يكون قلق الاختبار مفيداً من الناحية التربوية إذا أمكن توزيعه بمستوي منخفض نسبياً خلال مدة التعليم بدلاً من تركيزه عند مستوي مرتفع نسبياً قبل الامتحان مباشرة وأثناءه ، أن التعليم الجيد يتضمن تصريف الطاقة الناتجة عن قلق الاختبار بصورة تلقائية .

ويشير " جوهنز " في كتابة القياس والتقويم في الفصل المدرسي إلي أن المعلم يستطيع أن يخفف من القلق بأساليب عديدة قد تنفع الدارسين عندما يواجهون المواقف الانفعالية في مواقف أخري في حياتهم وهناك أساليب يمكن بواسطتها علاج القلق مثل التنفيس والفهم والتفكير الايجابي والاسترخاء .

ثانياً الحكمة الاختبارية :

كما ذكرنا قبل ذلك أن أداء الطلاب للامتحانات التحصيلية يتأثر ببعض العوامل الخاصة بالمفحوصين ومنها الحكمة الاختبارية أو المهارة الاختبارية التحصيلية الموضوعية فقد يكون ذلك نتيجة ضعف الحكمة الاختبارية لدي المفحوص يؤدي إلي انخفاض أدائه في الاختبارات التحصيلية .

وإذا لم يكن لدي التلميذ قدر مناسب من المهارة في كيفية الإجابة عن أسئلة الاختبار فقد يحصل علي درجة اقل من درجته الحقيقية عندما يعرض الإجابة بصورة رديئة . عرض الإجابة بصورة جيدة ليست مهارة معقدة تتطلب تخصص مرتفع ويصعب تحقيقها ، ويستطيع أي فرد اكتساب خبرة من الإجابة علي عدد من الاختبارات التي مرت به أن يدرك من خبرته الشخصية كيف يشرد عن الإجابة الخطأ وكيف لا يكون منبهاً بصورة جيدة للتعليمات ، وكيف يقع في خطأ تجاهل بعض القواعد الهامة .

ويذكر " فتحي عبد الحميد " أن أكبر دليل علي أهمية الحكمة الاختبارية يأتي من الأبحاث التي درست أساليب حل المشكلة فيما يتعلق بإجابات الطلاب لمفردات اختبار موضوعي منا " بلوم ، بردني " (1950) " آرل ، كونرل ، وونتمان "(1964) ، " فريش " (1965) ، " فتحي عبد الحميد " (1987) .

وقد بينت دراسة " مليان ، بربارا " (1982) أن الحكمة الاختبارية يمكن أن تدرب ، ولذلك هناك تطورات في أداء الاختبار خلال تكرار الاختبارات .

أ- مفهوم الحكمة الاختبارية :

يعرف " فتحي عبد الحميد " (1987) الحكمة الاختبارية بقوله :" هي مقدار إفادة المفحوص من تعليمات وخصائص الاختبار والموقف الاختباري في الحصول علي درجة عالية دون معرفة كاملة لمحتوي الموضوع الذي تم الاختبار فيه ".

ولعل اكثر من التعريفات شيوعاً تعريف " جيب " (19696) للحكمة الاختبارية علي انها " قدرة المفحوص علي الإجابة الصحيحة للمفردات ذات الاختيار من متعدد متضمنة أدله خارجية . من أجل الحصول علي درجة دون معرفة لمادة الموضوع الذي تم الاختبار فيه " .

كما يعرف " ميلر " (1965) أن الحكمة الاختبارية هي :

" المقدرة علي الإفادة من خصائص وصيغ الاختبار والموقف الاختباري للحصول علي أعلي الدرجات والحكمة تبعاً لهذا التعريف لا تعتمد علي معرفة متناول الاختبار بمادة الموضوع أو الدرس " .

ويري " فاروق عبد الفتاح " انه يطلق أحيانا علي المهارة في اخذ الاختبار الحكمة الاختبارية ويستطيع الطلاب ذوو الحكمة في الاختبار الحصول علي درجات طيبة سواء كانوا يعرفون كثيراً عن الموضوع أم لا . ولذا يري بعض المتخصصين في تصميم الاختبارات الموضوعية تعتبر أدوات قياس الحكمة في الاختبار أكثر منها أدوات قياس التحصيل الدراسي.

ب- مبادئ الحكمة الاختبارية :

لقد أفادت نتائج البحوث السابقة في تقديم مخطط تمهيدي لمبادئ الحكمة الاختبارية وتشمل الحكمة الاختبارية استراتيجيات متباينة ولكي ينجح المفحوص في التقدم أثناء الموقف الاختباري عليه أن يستخدم هذه الاستراتيجيات وهي : الاستخدام الفعال للوقت والتخمين ، وتجنب الخطأ والاستدلال الاستنباطي ومراعاة القصد أو الهدف ، واستخدام دليل او رمز .

وقد تضمن المخطط التمهيدي للحكمة الاختبارية الآتي :

أولاً : عناصر مستقلة عن مصمم وهدف الاختبار منها :

أ- إستراتيجية استخدام الوقت .

ب- إستراتيجية تجنب الخطأ .

ج- إستراتيجية التخمين ز

د- إستراتيجية الاستلال الاستنباطي .

ثانياً عناصر معتمدة علي الهدف من الاختبار وعلي مصمم الاختبار :

أ- إستراتيجية مراعاة القصد أو الهدف .

ب- إستراتيجية استخدام دليل أو رمز .

وقد قدم " مليمان بيشوب وابيل "(1965) وصفاً لمبادئ الحكمة الاختبارية وسوف نوضح ذلك فيما يلي :

أولاً : استراتيجيات استخدام الوقت : ويمكن تحقيقها عن طريق ما يلي :

1. البدء في العمل بسرعة مع تحقيق قدر معقول من الإتقان . أن السرعة في العمل مع تحقيق قدر معقول من الإتقان تتفاوت بين الأفراد . فقد يحاول فرد تكملة الاختبار في زمن أقل من الزمن المخصص للاختبار لكي يخصص زمنا للمراجعة .

2. وضع قائمة أو جدول لما يحدث أثناء الاختبار . وذلك لتحديد كم المطلوب من المفحوص عندما ينقص جزء معين من الاختبار ويتطلب الأمر قيام المفحوص بمراجعة دورية بانتظام من اجل الحفاظ علي السرعة المنابة لسيره في الاختبار .

3. إرجاء او تخمين المفردات التي لا يستطيع الفرد الإجابة عليها بسرعة فعندما يكون زمن الاختبار محدداً . يجب علي المفحوصين ان يجيبوا عن المفردات التي تنتج درجات اكثر في قدر اقل من الزمن المعطي للاختبار .

4. الاستفادة من الزمن المتبقي بعد تكملة الاختبار في المادة ويجب أن يقوم المفحوصين بتعديل الإجابات عند مراجعتها إذا كان ذلك يزيد من درجاتهم ، ولقد وجد أن المثابرة وهي تعني هنا استخدام كل الزمن في الاختبار عامل هام في زيادة درجات المفحوصين .

ثانياً إستراتيجية تجنب الخطأ :وتتمثل فيما يلي :

1. انتباه المفحوص للمفردات بعناية ، وتحديد طبيعة السؤال والمطلوب بوضوح.

2. علي المفحوص الا يخمن أو يستنتج الإجابة قبل قراءة السؤال بالكامل وعليه أن يتدرب بعناية علي إجابة المفردات التي لها أكثر من إجابة واحدة .

3. يجب الانتباه إلي خطورة الفشل في قراءة أو سماع تعليمات الاختبار تلقي أو تقرأ وأدراك خطورة الفشل في إتباعه بدقة .

ثالثاً : إستراتيجية التخمين :

احتمال الحصول علي درجة كبيرة في الاختبار الموضوعي الجيد عن طريق التخمين الاعمي فقد يكون ضئيلاً والتلاميذ المتفوقون الذين يتوفر لديهم الزمن الكاف لمحاولة الإجابة علي كل الأسئلة لا يلجئون إلي التخمين الاعمي إلا في حالات نادرة ويجب تشجيع التلاميذ علي التخمين العقلاني في إجابة فقرات الاختبارات الموضوعية التي لا يكونون متأكدون من صوابها تماماً .

وتتمثل إستراتيجية التخمين في الأتي :

1- قيام المفحوص دائماً بالتخمين إذا أعطيت الدرجة للإجابة الصحيحة فقط .

2- يخمن المفحوص لو أن تصحيح اثر التخمين اقل صرامة من معادلة أثر التخمين التي تعطي صفراً للإجابة العشوائية .

3- يخمن المفحوص لو كان التصحيح أو العقاب لاستخدام التخمين اكثر صرامة وذلك عندما تستبعد الاختبارات التي تعطي فرصة كافية للكسب والاستفادة .

رابعاً إستراتيجية الاستدلال الاستنباطي :

المفحوص الذي لديه حكمة اختباريه والذي لم يعرف الاختبار الصواب مباشرة من الممكن أن يكون قادراً علي أن يستدل أو يستنتج الإجابة بواسطة تحليل منطقي أو باستخدام معلومات مكتسبة من مفردات أخري . وهذه الإستراتيجية تختلف عن إستراتيجية دليل أو رمز . حيث فيها لم يكن من الضروري وجود ارتباطات بين الأدلة والإجابات الصحيحة . ومن الممكن استخدام إستراتيجية الاستدلال الاستنباطي بنجاح في أي موقف اختبار موضوعي ويعتمد نجاحها علي قدرة المفحوص علي أن يستدل بطريقة صادقة منطقياً .

خامساً : إستراتيجية مراعاة القصد أو الهدف :

قد يحصل المفحوص علي درجة منخفضة في اختبار موضوعي ، وذلك لأن أدائه واتجاهاته تختلف عن مصمم الاختبار ، وإذا استطاع المفحوص التعرف علي نزعات وأراء وتوجيهات مصمم الاختبار ومعرفة القصد في الاختبار والعمل وفقاً لهما فانه يتجنب فقد الدرجات الناشئة عن سوء الفهم والتفسير ، علي أن استخدام هذه الإستراتيجية يفترض أن الهدف أو الغاية للمفحوص حصوله علي اكبر عدد ممكن من الدرجات .وتتحقق هذه الإستراتيجية عن طريق ما يلي :

1. تفسير وإجابة الأسئلة وفقاً لأداء مصمم الاختبار أو الهدف من الاختبار .

2. إجابة المفردات حسب ما يقصد مصمم الاختبار .

3. توقع التعقيد أو التضليل .

سادساً : إستراتيجية استخدام دليل أو رمز :

الاستخدام الناجح للأدلة يعتمد علي الاتصال أو الاحتكاك السابق للمفحوص ، بالاختبارات المتشابهة ، وذلك يدعم العلاقة بين الأدلة والإجابة الصحيحة . وتستخدم الأدلة والرموز عندما يكون المفحوص غير قاد علي الأصول للإجابة بواسطة معرفته لمادة الموضوع وقدرته الاستدلالية .

وتتحقق هذه الإستراتيجية عن طريق :

1- التعرف علي الخصوصيات الثابتة لدي مصمم الاختبار والاستفادة منها في تمييز الإجابة الصحيحة من بين الاختبارات غير الصحيحة .

2- التعرف علي التحديدات الخاصة " الكلمات التي قد تكون مرتبطة بالإجابات الصحيحة أو غير الصحيحة " والاستفادة من ذلك عند إجابة المفردة .

3- التعرف علي التشابهات بين الاختبارات ، والاستفادة منها في إجابة المفردة .

4- الأخذ في الاعتبار مادة الموضوع وصعوبة المفردات المجاورة عند تفسير وإجابة المفردة.

ونجد مثلاً في بعض الأحيان أن بعض الطلاب يقومون بالإطلاع علي الامتحانات السابقة في مادة ما لمحاضر معين ، وعن طريق إجابتهم عن أسئلة هذه الامتحانات قد يتعرفوا علي فلسف هذا المحاضر في التقويم وبذلك يكتسبون خبرة تساعدهم علي الإجابة علي أسئلة الاختبار .

ويشير " فاروق عبد الفتاح " (1990) إلي بعض المهارات التي يجب توافرها عند الإجابة علي أسئلة الاختبار وهي :

1- يجب إدراك القاعدة التي سوف تعطي علي أساسها الدرجات أي هل ستخصم درجات الإجابات الخطأ من درجات الإجابات الصواب . هل تعطي درجات علي إجابات الأسئلة أكثر من الدرجات التي تعطي أعلي إجابات أسئلة أخري .

2- يجب معرفة أن الدرجة التي يمكن الحصول عليها سواء من المصحح أو من الآلات تعتمد إلي حد كبير علي وضوح الإجابة ودقتها . وعلي ذلك يجب بذل المزيد من الجهد في كتابة الإجابات أو تسجيل العلامات في ورقة الإجابة .

3- يجب أن يهيئ التلميذ نفسه جيداً قبل وقت الاختبار من الناحيتين الجسمية والعقلية فالتعب الناتج عن السهر طول الليل حتي مع استخدام المنبهات . يعتبر معوق ثقيل لجودة الأداء .

4- يجب الالتزام بالهدوء والروية حتي يتيسر الزمن الذي يستطيع فيه التلميذ الإجابة عن كل الأسئلة ، وهذا يقتضي عدم إضاعة وقت طويل في سؤال صعب أو مسألة معقدة أو الكتابة بإسهاب أكثر مما ينبغي في سؤال المقال .

5- يجب معرفة أن التحمن العادي غير ضار ولكن لا يجب أن يكون هو القاعدة في الإجابة ويجب أن يعمل التلميذ بأقصى جهد اثناء تأدية الامتحان بحيث يجيب علي كل الأسئلة حتي تلك التي لا يكون متأكد من صحة أجابتها تماماً .

6- يجب أن يقوم التلميذ بمراجعة إجاباته . إذا أمكنه ذلك وتيسر الوقت حتي يستطيع تصويب اي أخطاء غير مقصودة .. ويطلق في كثير من الأحيان علي هذه المهارات الحكمة في الاختبار .

اثر الحكمة الاختبارية علي اداء الاختبار التحصيلي :

تشير دراسات موثوق بها تأثير الحكمة الاختبارية على درجة الاختبار التحصيلي ، واوضح بعض المؤلفين مثل " ثورنديك " (1951) " ابيل ، دمرين " (1960) ، ابيل " (1965) ان الحكمة الاختبارية مصدر هام لتباين فر درجات الاختبار ، وابرز " ابيل " أن أكثر الخطأ في القياس قد ينشأ مما لدي الطلاب من مهارة تفضيل القليل جدا عند التعامل مع الاختبارات . وفي كلمات أخري يذكر " ابيل " أن المفحوص الذي لديه حكمة اختباريه عالية وذلك عند أداء الاختبار . وإذا كان كلام " أبيل " صادقا ، فانه يجب عمل محاولة لتعيين الطلاب الذين لديهم حكمة اختباريه منخفضة وعندئذ يجب أن يتعلم هؤلاء الطلاب الحكمة الاختبارية .

ولقد وجد كل من " مليان ، بريارا " (1982) في دراستهما بعنوان الحكمة الاختبارية كعامل في أداء اختبارات القراءات على الأطفال الأمريكان والمكسيكيين ، وكانت تهدف إلي معرفة آثار التدريب علي مهارات أخذ الاختبار علي أداء الأطفال في مرحلة الروضة – وجدوا أن معظم النتائج تدل علي أداء أفراد المجموعة المدرية أفضل من أداء المجموعة غير المدرية وتقترح أن الحكمة الاختيارية عامل يجب أن يؤخذ في الاعتبار عند تخطيط وإعداد برامج الاختبارات ووجدت الدراسة ايضاً أن الحكمة الاختبارية قد يكون لها أهمية خاصة عند اختيار أطفال من بيئات اقتصادية منخفضة .

ولقد وجد أن تعلم مبادئ الحكمة الاختبارية يؤدي إلي تحسين الأداء للاختبار التحصيلي وغالباً ما يحصل المفحوص الذي لديه حكمة اختباريه علي درجات أعلي من المفحوص المتساوي معه في القدرة العقلية وذلك في الاختبار التحصيلي ، وعلاوة علي ذلك فأنه يختلف معني الافتراض أن القدرة علي استعمال مبادئ الحكمة الاختبارية ذات اتصال علي بدرجة المعرفة لمحتوي الموضوع . أن كثيرا من التلاميذ يحضرون إلي المدرسة وهم في حاجة إلي الخبرة فيما

يتعلق بطرق التعامل مع الاختبارات وقد يكون ضعف أدائهم ليس بسبب ضعف قدرتهم المعرفية ، وإنما بسبب معرفتهم غير الكافية عن كيفية التعامل مع الاختبارات .

قياس الحكمة الاختبارية :

لقد استفادت الدراسات التي أجريت في هذا المجال مما قدمه " مليمان وابيل " (1965) من وصف وتحليل الحكمة الاختبارية ولقد تبنت بعض الدراسات استراتيجيات الحكمة الاختبارية في القياس واقتصر البعض الآخر من الدراسات في قياس الحكمة الاختبارية علي بعض من هذه الاستراتيجيات .

ولقد تبني " فتحي عبد الحميد " إطار " مليمان ، بيشوب وابيل " لاستراتيجيات الحكمة الاختبارية الست ، وقام بإعداد مقياس للحكمة الاختبارية متضمنا كل الاستراتيجيات التي وجدت في إطار وصف وتحليل " مليمان ، بيشوب وأبيل " (1965) للحكمة الاختبارية .

وفي النهاية ومن أجل تحقيق الإمكانية التعليمية للاختبارات كعينات تعليمية ، ومن الواجب أن يؤخذ الاختبار بعين الاعتبار في أثناء التخطيط للتعليم كما يجب أن يلعب دورا ذو دلالة في مختلف مراحل التعليم من البداية إلي النهاية ويجب أن يؤخذ في الاعتبار جميع العوامل التي تؤثر في أداء الاختبار .

ثالثاً : التدريب

أ – انماط التدريب :

1- المران Coaching :

النمط الأول للتدريب الذي ظهر بصورة واسعة هو ما يسمي بالمران والذي استخدم بصورة كاملة في برامج التدريب وكان واضحا أن البرامج التي اشتملت علي هذا النمط اشتركت جميعها في التشابه بين محتوي الاختبار ومادة المران . وأن برنامج المران اقتصر في الغالب علي محتوي معين ارتبط سواء بصورة مطابقة أو بصور مماثلة بما يدور حوله الاختبار وما يحتويه من مفردات وأنه يعتبر وسيلة الهدف الرئيسي منها زيادة الدرجة في الاختبار وتم في هذا النمط شرح وتحليل طرق الإجابة عن مفردات متشابهة أو مطابقة لمفردات الاختبار مع التدريب علي أسئلة تدريبية من نوع مفردات الاختبار المستخدم ، وعرض المحتويات الأساسية بالاختبار ونع المفردات والمراجعة للأسئلة والتمارين والتعريف بالاختبار والهدف منه وطريقة تقدير الدرجات ومفتاح التصحيح ، والقيام بتطبيق اختبارات تشمل عينة من مفردات الاختبار أو مفردات مماثلة لها يقوم الفرد بالاطلاع عليها وإجابات ، ومزودة بالإجابات الصحيحة للأسئلة والذي اعتبر تغذية مرتدة . واستخدمت المحاضرات والمناقشات ونسخ الاختبارات المتشابهة ونشرات للمعلومات وظهر هذا النمط في معظم الدراسات التي تم عرضها تحت محتوي برامج ودورات التدريب .

وقد يكون أقرب مفهوم للمران هو ما يسمي بالتدريس للاختبار حيث يقوم المعلم بتدريب التلاميذ وإعدادهم للإجابة عن أسئلة معينة يتضمنها الاختبار ومحاولة تثبيت هذه الإجابات في أذهان التلاميذ من أجل يوم الامتحان ، بمعني أن التدريس يقوم علي تغطية محتوي أسئلة الاختبار — وينفق هذا مع ما يقوم به الطلاب من توافق طرق استذكارهم طبقاً لنوع الاختبار الذي يطبق عليهم ، ولذا فقد لا يهتمون باستخدام الكتب والمذكرات لاستيعاب الأفكار لأنهم يعرفون أن هذه الأفكار لا تساعد كثيراً في حل أسئلة الاختبارات الموضوعية وكثيراً ما نجد أسئلة بالكتب مشابهة للأسئلة الموجودة بالاختبار . ووفقاً لما ذكره

Jacobs,p(1966) فان مصطلح المران يستخدم على انه تدريس مركز بصورة جماعية أو فردية ويتم في أماكن خارج المدرسة ، ويوجه مباشرة إلي زيادة الدرجة في الاختبار بصورة مرتفعة بغض النظر عن التعليم الجيد .

ويذكر (Jackson , R (1980 ان كلمة مران تشير إلي طرق عديدة تستخدم كمحاولة لزيادة كل من مهارات التلاميذ – الذين لديهم وقت قصير نسبياً – المرتبطة بمحتوي الاختبار، وقدرات الاستدلال التي يقيسها بالاختبار، وهذا المعني يطابق التدريس ، ويؤكد علي أن المران يعني بالضبط حشو عقول التلاميذ بعينة من مفردات الاختبار في فترة وجيزة من الوقت .

ويشير (Vernon, p (1954 إلي أن المران يعطي الفرد بالإجابات الصحيحة لأسئلة الاختبار والساعات عن كيف يحسن أداءه في الاختبار ويتم في انجلترا إعطاء مران في المدارس علي صور متشابهة تماما للاختبارات المستخدمة في الانتقاء .

ويذكر(Cronbach , L (1970 أن المران علي اختبارات الرياضيات واللغة احتوي علي تدريب إضافي مركز وشمل المران علي اختبارات القدرة العقلية العامة دراسة الاختبارات المستخدمة في العام الماضي والشئون علي مشكلات الاستدلال الموجودة في الاختبارات وكيفية الإجابة عن مشكلات الاختبار بسرعة .

ويذكر(petty & porter , (1980 أن المران يعني إعطاء تعليم لمشكلات سائلة لمفردات الاختبار وقد وصف مفهوم المران بأنه عملية شحن الفرد بأكبر قدر ممكن من مشكلات الاختبار وحلولها ، حيث اورد (Slack & porter (1980 ان المران اشتمل علي تدريب لمفردات اللغة وحقائق الحفظ والاستظهار وما شابه ذلك من تمرينات معينة للإعداد علي مفردات مماثلة للموجودة في الاختبار

وقامت هيئة الجامعة بأمريكا بمسح برامج المران في سبع ولايات في شمال شرق امريكا واتضح ان الطرق المتبعة في معظم برامج المران تحتوي علي تدريب مكثف لعينة من أسئلة الاختبار ، وأن المران هو تدريب مكثف علي مفردات مشابهة لتلك الموجودة في الاختبار .

وتضيف (1981) Anastasi ان المران يطور من مهارات خاصة تكون اقل استخداماً في انشطة الحياة اليومية ، وان هذا هو ما يحدث عند التدريس للاختبار حيث يميل الفرد إلي التركيز علي عينة من المهارات والمعلومات الخاصة التي تغطي الاختبار بالإضافة إلي النواحي المعرفية التي يغطيها الاختبار والتعرف علي الأشياء التي تخيفها عن المفحوص للحفاظ علي صدق الاختبار والتركيز علي المهارات الأساسية في الاختبارات ذات التطبيقات المحددة .

ويري (1981) Hopkings & stailty) ان المران يقصد به اي شئ من التدريب للمفردات التي حذفت Missed من الاختبارات وتستخدم في برامج المران العامة .

ويري (1981) Messkick & jongeblut ان البعض يقصد بالمران تمرن لعينة من مفردات الاختبار والمراجعة النهائية لدخول الامتحان .

ويري Meherens & Lehmann ان المران يقصد به عادة جلسات للتعليم خاصة وقصيرة الأجل وتشمل التدريب علي عينة من مفردات مشابهة لتلك الموجودة في الاختبار .

ويتضح مما سبق أن استخدمت بعض المرادفات لمصطلح المران مثل : التدريب ، والحشو ، والتمرن المركز علي مفردات هي نفسها او مشابهة للموجودة في الاختبار ، وانه في حالات كثيرة يتطابق في مفهومه مع مفهوم التدريس للاختبار وقد ياتي ذلك مناقضا لما ذكره (1950) Guilford ان مفهوم

التدريب اصطبغ – بدون وجه حق – بصبغة جعلته يشبه بتقوية العضلات بالتمرن ، وأن التدريب بمعناه النفسي لا يمكن أن ينظر إليه بهذه الكيفية .

أن المران يقصد به شكل من الممارسة المقصودة لأسلوب الاختبار ونوع بنوده وطريقة الحصول علي الدرجة المقبولة فيه من خلال التعرف علي أساس الإجابة وفهم المقصود من الاختبار بمعني أن المفحوص تدرب علي كشف الجوانب الفنية في الاختبار التي تسعي الأصول العلمية جعلها خالية من المفحوص للاحتفاظ بصدق الاختبار ، بالإضافة إلي دراسة (1936) انه عندما لا تحدث الخبرة زيادة في الذكاء نفسه فأنها تحدث الزيادة في درجات اختبارات الذكاء حيث يكون التشبع بالعامل العام علي وجه العموم .

ويشير (Hopking & Stanley (1981 إلي ان أغلب التلاميذ يظهرون – بصورة منسقة – تحسنا عاما في الدرجة عند أعادة الاختبار وأن متوسط الزيادة بلغ (2.5) نقطة في نسبة الذكاء عندما اعطيت صورا متكافئة للاختبار .

كما وجدت (Kreit (1968 له (7) نقاط عندما تم تطبيق أربعة اختبارات ذكاء لتلاميذ الصف الثالث بعد فترة مدتها خمس شهور .

ويذكر (Rutan (1979 ان الأفراد الذين ليس لديهم الفة بالاختبارات الموضوعية أو اختبارات المقال يكون اداؤهم عادة ضعيفا – إلي حد ما – بالمقارنة بأداء الأفراد الذين اعطيت لهم ثلاثة اختبارات جمعية شائعة للمرة الثانية أو الثالثة – في فترات بلغت العام – حصلوا علي درجات أعلي من الأطفال الذين في نفس السنة الدراسية وأعطيت لهم الاختبارات للمرة الأولي . وتوجد إشارات كثيرة تدل علي ان اعادة الاختبار تؤدي إلي زيادة الأداء في مفردات الاختبار . وحتى في " الكونغو " وجد أن نتائج التطبيق الأول غير ثابتة إلي حدا ما لدي

الأفراد الذين ليس لديهم خبرة بأخذ الاختبار ، كما وجدت زيادة في درجات اختبار المصفوفات المتتابعة عندما طبق مرتين وثلاث مرات .

ويتضح مما سبق ان الخبرة لها تأثير موجب علي الأداء ، إلا ان مسألة تحديد هذا الأثر قد تكون مضللة أن لم نتناوله في ضوء الكثير من المتغيرات التي تتدخل عند إعادة التطبيق ، ولعل التساؤل الذي أورده (1932) (Terman) ومؤداه : هل ستعطي اختبارات الذكاء نفس النتائج عندما تطبق للمرة الثانية ؟ يوضح ذلك ؟ ويجيب بأنه يمكن أن يحدث ذلك بشرط أن نضع في اعتبارنا بعض المتغيرات مثل الذكاء والفترة بين التطبيقات ، هذا بالإضافة إلي متغيرات أخري مثل تطبيق نفس الاختبار أو تطبيق صور متوازنة ، والعوامل الثقافية وطبيعة الاختبارات ، فقد ذكرت " أنستازي " (1976) أن الكثير من الدراسات تناولت أثر تكرار تطبيق اختبارات الذكاء علي مدي فترات تمتد من أيام قليلة إلي سنوات عديدة وضمت المراهقين والأطفال والأشخاص الأسوياء والمتأخرين عقلياً ، وشملت اختبارات فردية واختبارات جمعية ، وأجمعت الدراسات علي وجود زيادات دالة إحصائياً عند إعادة تطبيق الاختبار ، وأن هذا التحسن لا يكون محددا ضرورياً للتكرارات الأولية ، وأنه عندما يوجد استمرار للمكاسب فأن ذلك يشير إلي الاعتماد علي صعوبة الاختبار ومستوي القدرة .

وتضيف " انستازي " (1981) أن درجة الاختبار في التطبيق الثاني تميل إلي أن تكون اكبر منها في التطبيق الأول وانه وجد متوسط زيادة دال إحصائياً عند تطبيق الاختبار باستخدام صور تبادلية لمرات أخري بفاصل زمني يبدأ من يوم ويمتد إلي ثلاث سنوات ونفس النتائج تم التوصل إليها مع الطلبة المتفوقين عقلياً ودراسياً ولدي طلبة المدارس الثانوية والجامعية وعينات من الموظفين في حالة استخدام اختبارات متوازنة أيضا " ولعل تلك الاستنتاجات تجعلنا نعرض أثر الخبرة علي الأداء في ضوء مجموعة من المحددات .

رابعاً : التخمين

يستطيع بعض الطلبة زيادة درجاتهم في الاختبارات الموضوعية عن طريق التخمين ، ويلجأ المتعلم للتخمين لعدم وجود المعلومات الكافية أو القدرة علي الإجابة . ويجد أكثر من نوع للتخمين ، فهناك التخمين الأعمى — وفيه يختار المفحوص الإجابة عشوائياً دون تدقيق في محتواها ، ويوجد التخمين المبني علي معلومات (البناء) ، وفيه يدقق المفحوص في مقدمة السؤال والاختبارات المطروحة ومحتواها والفرق بينها ومدي ارتباط وعلاقة كل اختبار منها بالسؤال حتي يصل للإجابة الصحيحة ، وفي هذه الحالة يكون لدي المفحوص الحافز في الوصول للإجابة الصحيحة ومعلومات جزئية عن السؤال .

ويظهر التخمين في أسئلة الصواب والخطأ وأسئلة الاختيار من متعدد ، ولكن كلما زادت الاختبارات كلما زادت الاختيارات كلما قل أثر التخمين ، ويستخدم المفحوص التخمين البناء إذا كان ليس لديه بعض المعلومات الجزئية ، ويستخدم التخمين الأعمى إذا كان ليس لديه أي معلومات علي الإطلاق أو معلوماته ضعيفة جداً وغير واضحة ولا متكاملة .

يري البعض أن التخمين مشكلة ويجب أن تعاقب المفحوص الذي يلجأ للتخمين خصوصاً التخمين الاعمي ، ولابد من التعامل مع هذه المشكلة بتفكير كافي وعن طريق الوسائل التي يستخدمها عند إعداد الاختبار ، وبذلك نستخدم الدرجات المعيارية للاختبار . وبالرغم من أن هذا كان مصدر خلاف لعدة سنوات ، والآن أصبح الدليل واضح ضد استخدام التخمين ، وقبل أن ننظر لهذا الدليل لابد أن ندرك أننا في حاجة لتوجيه التخمين والأسباب التي دعتنا إلي عدم تشجيعه هي :

السبب الأول : الاعتقاد الأخلاقي بخطأ التخمين لأنه يعتبر شكل من أشكال المغامرة وشد الانتباه للخداع .

السبب الثاني : أن التخمين يؤثر علي سمات الاختبار . وتوجد صيغة لإلغاء تأثير التخمين الأعمى في معظم أشكاله وهي :

الدرجة الصحيحة = عدد الإجابات الصحيحة – عدد الإجابات الخاطئة

عدد البدائل – 1

خامساً: طرق الاستجابة

سوف نستخدم طرق الاستجابة لكي نشير إلي السمات الشخصية أو القدرة وطريقة تأثيرها علي استجابات الفرد في الاختبار . فأسلوب استجابة الطالب الذي لديه حكمة اختباريه ، يعرف كيفية التعامل مع مفردات الاختبار ، فإذا أدركها وطبقها فأنها سوف تشكل أسلوبا للاستجابة يميزه عن غيره من الطلاب . والميل إلي التخمين للاستجابة يميزه عن غيره من الطلاب .

والميل إلي التخمين يعتبر أسلوب استجابة يستخدمه الطالب إذا كان لا يعرف أو غير متأكد من الإجابة الصحيحة علي السؤال . فالحكمة الاختيارية والتخمين لهما تأثير علي درجة الطالب في الاختبار كما أن السرعة في العمل أو الإجابة علي حساب الدقة لها تأثير علي درجة الطالب في الاختبار . وأما أساليب الاستجابة الموضوعية فأنه يمكن التحكم فيها من خلال التعليمات الجيدة للاختبار .

سادساً : ترتيب مفردات الاختبار

لا يزال المهتمون بالقياس يعتبرون أن ترتيب مفردات الاختبار متغير هام يؤخذ في الاعتبار عند إعداد الاختبارات . وقدمت اقتراحات لترتيب المفردات منها ، الترتيب وفقاً لنوع المفردة ، الترتيب حسب صعوبة المفردة ، الترتيب وفقاً لنوع المفردة ، الترتيب حسب صعوبة المفردة ، الترتيب وفقاً للموضوعات ذات الصلة ، والترتيب

وفقاً لعرض المادة الدراسية في الفصل . ولقد نصح خبراء القياس المعلمين بترتيب المفردات من السهل إلي الصعب أي وفقاً لتدرج الصعوبة ، وذلك من أجل أسباب دافعية أن الطلبة الذين تلقوا مفردات أسهل في بداية الاختبار التحصيلي استطاعوا أن يعدلوا من قلقهم بالنسبة للاختبار بسرعة كبيرة ، ومن ثم ازداد اداؤهم ، وبالعكس إذا كانت المفردات مرتبة من الصعب إلي السهل تؤدي إلي خفض الداء ، وذلك يرجع إلي الأثر الناتج عن وضع المفردات الصعبة في بداية الاختبار . فعند مقارنة أداء الطالب في حالة ترتيب المفردات من السهل إلي الصعب بترتيبها من الصعب إلي السهل ، أو العشوائي ، فإن الفرق في الأداء يعزي إلي التكيف والتخفيف من القلق الاختباري في حالة الترتيب من السهل إلي الصعب ، أو يعزي إلي الأثر في حالة الترتيب من الصعب إلي السهل أو يعزي إلي كليهما معاً . ووجد أن ترتيب المفردة له تأثير دال علي أداء الاختبار ، ولكن هذه العلاقة ظهرت عندما أجريت الاختبارات كاختبارات سرعة أكثر من اختبارات القوة .

ولقد وجد أن تقديم معلومات للمفحوصين عن صعوبة المفردات يحسن أدائهم ويساعدهم في الحصول علي درجات أعلي في الاختبارات خاصة الاختبارات التحصيلية .

وأخيراً : كيفية قياس الحكمة الاختبارية :

وفي سبيل إعداد مقياس للحكمة الاختبارية تم أجراء الخطوات التالية :

1. الدراسة الاستطلاعية .
2. إعداد المقياس في صورته الأولية .
3. التجريب المبدئي للمقياس .

4. حساب ثبات وصدق مقياس الحكمة الاختبارية .

5. التحليل العاملي لمقياس الحكمة الاختبارية .

6. إعداد مقياس الحكمة الاختبارية في صورته النهائية

7. الزمن المناسب لمقياس الحكمة الاختبارية .

8. المعايير الخاصة بمقياس الحكمة الاختبارية .

أولا: الدراسة الاستطلاعية :

في ضوء ما أسفرت عنه بعض الدراسات والبحوث الأجنبية حول موضوع الحكمة الاختبارية (T .W) ، فقد تم الحصول علي بعض العناصر المتعلقة بالحكمة الاختبارية والناتجة من استفتاء مفتوح قدم لطلاب المدرسة العليا بالولايات المتحدة حول المساعدات والتوجيهات التي تعطي للمفحوص ، وتفيده في الموقف الاختباري عند أدائه للاختبار في أي مادة دراسية . وتم رصد هذه العناصر فيما يتعلق بالحكمة الاختبارية فيما يلي:

م	العناصر	الإستراتيجية
1	اقرا الأسئلة بحرص وعناية	تجنب الخطأ
2	لا تستغرق وقتاً طويلاً في إجابة سؤال واحد	استخدام الوقت
3	راجع إجاباتك من الخطأ	تجنب الخطأ
4	خمن إذا لم تعرف الإجابة الصحيحة	التخمين
5	استبعد الاختيارات التي من المعروف أنها غير صحيحة	الاستدلال الاستنباطي
6	استفد من أي إيحاءات من الأسئلة الأخرى في	الاستدلال الاستنباطي

	الإجابة	
7	أجب الأسئلة السهلة أولاً ثم الأسئلة الصعبة	استخدام الوقت
8	استغل الوقت جيداً	استخدام الوقت
9	انتبه للتحديدات الخاصة مثل كلمة : أبداً – دائماً – نادراً – جميعاً	استخدام دليل
10	راع قصد السؤال وهدفه	مرعاه القصد

ويتضح من إجابة الطلاب بالولايات المتحدة علي الاستفتاء المفتوح ، والمتعلق بالحكمة الاختبارية .

أن جميع العناصر التي ذكرت تتبع استراتيجيات الحكمة الاختبارية الست (استخدام الوقت – تجنب الخطأ – التخمين – الاستدلال الاستنباطي – مراعاة القصد أو الهدف – استخدام دليل).

وكذلك فقد استعين بعض المقاييس السابقة للحكمة الاختبارية لكل من " سلكتر وآخرين " (1970) " جيمس ديموند ، ايفانز" (1972)، " كيفين كريهان وآخرين " (1974) ، جيمس ديموند وآخرين " (1976) ، " جون باجتلست " (1977) " وتم رصد الاستراتيجيات وتكرارها من المقاييس السابقة ، ويتضح ذلك من الجدول التالي :-

جدول رقم (1)

استراتيجيات الحكمة الاختبارية وتكرارها ونسبتها المئوية من المقاييس السابقة

الإستراتيجية	التكرار	النسبة
الاستدلال الاستنباطي	3	0.375
استخدام دليل او رمز	5	0.625

وقد تم تصميم استفتاء مفتوح للتطبيق علي عينة من طلاب الصف الأول الثانوي مكونة من (78) طالباً وطالبة (فصل من البنين ، وفصل من البنات) ، حيث طلب فيه من كل منهم أن يذكر الأمور التي يجب علي المفحوص مراعاتها في الموقف الاختياري وعدد أدائه للامتحان، والتي تمكنه من الحصول علي درجة عالية فيه ، دون معرفة كاملة بمادة موضوع الامتحان.

وقد ركز الطلاب في إجاباتهم علي الأمور الثانية (تكررت من جانب 80 ٪ من الطلاب):

1. العناية بقراءة الأسئلة جيداً وفهم المطلوب منها .
2. قراءة التعليمات بعناية أو وجدت .
3. إجابة الأسئلة السهلة والتي لا تستغرق وقتاً طويلاً أولاً ثم التفكير في إجابة الأسئلة الصعبة بعدها
4. تخصيص وقت محدد لإجابة كل سؤال حسب ما يخصه من زمن الامتحان .
5. توفير جزء من الوقت لمراجعة الإجابات .
6. تجنب السرعة والتأني في الإجابة .
7. التخمين خاصة عندما يكون وقت الامتحان غير كاف أو الأسئلة صعبة .
8. إجابة جميع الأسئلة وعدم ترك أي سؤال بدون إجابة .
9. الاستدلال علي إجابة السؤال من إجابة سؤال آخر .
10. الاهتمام بمعرفة قصد السؤال ، أو واضع السؤال (الممتحن) قبل الإجابة .

11. في حالة أسئلة الاختيار من متعدد ، تستبعد الاختيارات التي من المعروف لدي المفحوص أنها غير صحيحة والاختيار من بين البدائل الاختيارية المتبقية .

12. مراجعة الإجابات في نهاية الامتحان وقبل تسليم ورقة الإجابة .

13. في أسئلة الاختيار من متعدد ، تراعي التشابهات بين الاختيارات عند اختيار الإجابة .

ويفحص العناصر التي نتجت من إجابة الطلاب للاستفتاء المفتوح وجد أنها تنتمي إلي استراتيجيات الحكمة الاختيارية ، والتي حددها " مليمان وأبيل " (1965) .

حيث وجد أن العناصر (1 ، 2 ، 12) تنتمي إلي إستراتيجية تجنب الخطأ ، والعناصر (3 ، 4 ، 5 ، 6) تنتمي إلي إستراتيجية استخدام الوقت ، والعنصرين (9 ، 11) ينتميان إلي إستراتيجية الاستدلال الاستنباطي ، والعنصر (10) ينتمي إلي إستراتيجية مراعاة القصد أو الهدف ، والعنصر(13) ينتمي إلي إستراتيجية استخدام دليل أو رمز .

وبذلك تم الحصول علي ثلاثة أنواع من تكرارات الاستراتيجيات الخاصة بالحكمة الاختبارية أحدها من الإطار النظري والدراسات السابقة ، والثاني من المقاييس السابقة ، والثالث من الاستفتاء المفتوح ، وعلي أساس متوسط التكرار لكل إستراتيجية من استراتيجيات الحكمة الاختبارية الناتج من الأنواع الثلاثة المذكورة سابقاً ، أمكن تحديد الأهمية النسبية لكل إستراتيجية من استراتيجيات الحكمة الاختبارية ، وفيما يلي استراتيجيات الحكمة الاختبارية مرتبة وفقاً لأهميتها في قياس الحكمة الاختبارية :

1. إستراتيجية استخدام دليل أو رمز .

2. إستراتيجية الاستدلال الاستنباطي .
3. إستراتيجية استخدام الوقت .
4. إستراتيجية تجنب الخطأ .
5. إستراتيجية التخمين .
6. إستراتيجية مراعاة القصد أو الهدف .

ثانياً : إعداد المقياس في صورته الأولية :

للوصول إلي الصورة الأولية للمقياس أجريت العمليات الآتية :

1. تحديد الأوزان النسبية للاستراتيجيات التي تتضمنها الحكمة الاختبارية :

تتطلب عملية بناء المقياس تحليل ميدان القياس وتقسيمه إلي عناصر أو مواضيع والكشف عن عدد أجزاء كل موضوع والأهمية النسبية لكل جزء ويتم اختيار نوع من المفردات يناسب قياس ذلك الميدان ، وصياغة موضوعات ذلك الميدان في أسئلة تمثل في مادتها وعددها ميدان القياس تمثيلا إحصائيا صحيحا

وقد تم حساب الوزن النسبي لكل إستراتيجية من الاستراتيجيات التي تتضمنها الحكمة الاختبارية في ضوء ما أسفرت عنه تكرارات هذه الاستراتيجيات في الإطار النظري والدراسات السابقة والمقاييس السابقة والاستفتاء المفتوح ، ويوضح ذلك الجدول التالي :

جدول رقم (2)

استراتيجيات الحكمة الاختبارية واهميتها النسبية .

م	الإستراتيجية	النسبية
1	استخدام دليل أو رمز	0.38
2	الاستدلال الاستنباطي	0.26
3	استخدام الوقت	0.12
4	تجنب الخطأ	0.08
5	التخمين	0.08
6	مراعاة القصد أو الهدف	0.08

2. صياغة مفردات المقياس :

وضع الباحث عددا من الأسئلة عن كل إستراتيجية يتناسب مع الأهمية النسبية للإستراتيجية وحدد لكل سؤال أربع إجابات (أ ، ب ، ج ، د) وعلي المفحوص اختيار إجابة واحدة فقط لكل سؤال ، واعتمد الباحث في وضع أسئلة وإجابات مقياس الحكمة الاختيارية علي ما جاء من أمثلة بالدراسات السابقة والمقاييس الأجنبية مع تعريب هذه الأسئلة وإجاباتها لكي تصبح صالحة للبيئة العربية ، بالإضافة إلي ما جاء في الاستفتاء المفتوح والإطار النظري . بعد صيغت مفردات (أسئلة) المقياس ، وتتكون من (38) سؤالا . لقياس الحكمة الاختبارية لدي المفحوص . حيث احتوت ، إستراتيجية استخدام دليل أو رمز علي

(15) سؤالا ، وإستراتيجية استخدام الوقت علي أربعة أسئلة ، وكل من إستراتيجية تجنب الخطأ والتخمين

ومراعاة القصد أو الهدف علي ثلاثة أسئلة . وقد روعي في صياغة المفردات الدقة ، ذلك أن النتائج قد تتغير طبقاً لصيغة الأسئلة والطريقة التي يطلب بها الإجابة عنها ، وأن تكون الأسئلة والإجابات في حدود فهم أفراد العينة وتتمن كل سؤال فكرة واحدة .

3. العرض علي المحكمين :

عرض المقياس علي لجنة من الحكمين تتكون من ثمانية من المتخصصين في التربية وعلم النفس التعليمي والصحة النفسية .

حيث قد للمحكمين التعريف الإجرائي للحكمة الاختبارية بالإضافة إلي التعريف الإجرائي لكل إستراتيجية من الاستراتيجيات المستخدمة في قياس الحكمة الاختبارية مع وجود تعليمات توضح لهم ماهية واستخدام المقياس ، وطبيعة العينة . وطلب منهم قراءة المفردات التي تتضمنها كل إستراتيجية ، وإبداء الرأي نحو انتماء كل مفردة للإستراتيجية التي تتبعها من عدمه .

وقد قام الباحث بإجراء بعض التعديلات في صياغة الأسئلة والإجابات بما يتناسب مع وجهة نظر المحكمين ، كما استبعدت الأسئلة التي حصلت علي نسبة اتفاق أو قبول اقل من 80 ٪ .

4. تحديد نوع الاستجابة :

روعيت ملاحظات المحكمين في الصورة للمقياس وتتكون من (36) سؤالاً ، وهذه الصورة تتكون من كراسة للأسئلة التي تبدأ بتعليمات ، يليها المقياس ، وهي من نوع الاختيار من متعدد ولكل سؤال أربع إجابات رموزها (ا ، ب ، ج ، د) ، ثم

ورقة منفصلة للإجابة بها أرقام الأسئلة ، وأمام كل رقم مربع ، علي المفحوص اختيار إجابة واحدة من الأربعة اختيارات لكل سؤال ، مع وضع رمز الإجابة التي يختارها في المربع الموجود علي يسار رقم السؤال .

5. تحديد التعليمات :

صيغت تعليمات المقياس بما يتناسب مع نوع مفرداته وتوفرات الدقة مع الوضوح في صياغتها حتي لا تختلف النتائج باختلاف فهم المفحوصين للتعليمات ، ولذلك تم تحديد المطلوب إجراؤه من المفحوص مع تقديم الهدف من استخدام المقياس ، والتأكيد علي أن نتائج هذه المقياس خاصة بالبحث العلمي وأن يطلع عليها احد سوي الباحث . وتم توضيح طريقة الإجابة علي المقياس في ورقة الإجابة المنفصلة التي يدون عليها المفحوص بيانات الاسم والمدرسة والصف الدراسي والجنس والعمر والزمن المستغرق في الإجابة

6. تقدير الدرجة :

في تقدير الدرجة اعتمد الباحث علي ثلاثة مصادر هي :

أ- خبرة الباحث لما تعرض له من دراسات وبحوث سابقة ومقاييس متعلقة بالحكمة الاختبارية .

ب- أوزان الإجابات من بعض المقاييس التي بنيت لقياس الحكمة الاختبارية لدي الطلاب .

ت- عرضت الأسئلة وإجاباتها علي مجموعة من الحكمين المتخصصين في علم النفس التعليمي والصحة النفسية ، طلب منهم اعطاء أوزان لإجابات كل سؤال تنفق مع درجة صحتها ، وقد روعي أن تكون مجموعة المحكمين هذه من

لجنة المحكمين التي عرض عليها المقياس وحتى يتوافر لديهم معلومات عن القياس تمكنهم من تحديد الأوزان للإجابات .

وفي سبيل ذلك ، فقد كانت الوزان لإجابات بعض الأسئلة علي الصورة (صفر ، 1) ، والبعض الآخر علي الصورة (صفر ، 1 ، 2) ، وللبعض الثالث (صفر ، 1 ، 2 ، 3) ، علي أنه توجد إجابة خطأ علي الأقل ضمن إجابات السؤال ، وفي إستراتيجيتي الاستدلال الاستنباطي واستخدام دليل أو رمز ، كانت إجابات السؤال تأخذ الوزن (صفر ، 1) اي انه توجد إجابة واحدة صحيحة وباقي الاختيارات خطأ ، أما إجابات كل سؤال في الاستراتيجيات الأخرى فقد أخذ أحد الأوزان السابقة .

ثالثا : التجريب المبدئي للمقياس :

بعد ذلك وزعت الأسئلة داخل المقياس عشوائياً ، كما وزعت الإجابات الأربع لكل سؤال توزيعاً عشوائياً ، وحتى لا يكشف المفحوص دلالة الإجابات في المقياس . وطبق المقياس في صورته المبدئية علي عينة قوامها (165) طالباً (ثلاثة فصول من البنين وفصلين من البنات) من طلاب الصف الأول الثانوي في شهر ديسمبر (1985) ، وكما هو موضح في الجدول رقم (3) .

جدول رقم 3:

العدد الكلى للعينة الاستطلاعية وعدد الطلاب

فى كل مدرسة من المدارس الممثلة للعينة الاستطلاعية .

المدرسة	الإدارة التعليمية	عدد الطلاب
أحمد عرابي الثانوية بنين	الزقازيق	31
الثانوية بنات	الزقازيق	34
جمال عبد الناصر الثانوية العسكرية	منيا القمح	32
الثانوية بنات	منيا القمح	36
بلبيس الثانوية العسكرية للبنين	بلبيس	32

ولقد راعى الباحث فى تطبيق المقياس عدم الرد على أسئلة المفحوصين إلا فى حدود تعليمات المقياس،

وذلك حتى لا يتجه المفحوص نحو إجابة معينة للسؤال وذلك حتى لا يتجه المفحوص نحو إجابة معينة

للسؤال . وطلب من المفحوص اختيار إجابة واحدة فقط لكل سؤال من أربع إجابات (ا،ب،ج،د) كما طلب من كل مفحوص تسجيل الزمن الذي استغرقه فى إجابة أسئلة المقياس كاملة وذلك بعد الانتهاء من إجابة المقياس ،وحتى يستطيع الباحث الإفادة من ذلك فى تحديد الزمن اللازم لإجابة المقياس فى ضوء متوسط الزمن للعينة الاستطلاعية . ولقد تم تصحيح المقياس باستخدام مفتاح تصحيح مثقب .

قدرة المفردات على التمييز :

للتمييز بين الإجابات الربع المختلفة ، ومدى توزيع هذه الاستجابات على الاحتمالات الاختيارية الأربعة لكل مفردة استخدام الباحث اختبار (كا^2). فعندما يكون للسؤال أربع إجابات محتملة (ن=4)،فإن التكرار المتوقع يساوى خارج قسمة المجموع التكراري على (4)، ودرجات الحرية فى هذه الحالة هى (3) . وبما أن حدود الدلالة الإحصائية لدرجات حرية (3) عند مستوى (0.05)= 7.835 ، وعند مستوى (0.01)= 11.35، وعند مستوى (0.001)= 16.27

وتحسب كا^2=مج (ت$_{و}$ – ت$_{م}$)

ت$_{م}$

حيث يدل الرمز على المجموع ،ت$_{و}$ على التكرار الواقعي ،ت$_{م}$ على التكرار المتوقع . ويبين الجدول رقم (4) حساب قيم كا^2 ومستوى دلالتها .

جدول رقم (4):

قيم كا^2 ومستوى دلالتها

رقم السؤال	كا2	رقم السؤال	كا2	رقم السؤال	كا2
1	223.24	13	90.83	25	72.40
2	425.76	14	321.01	26	365.10
3	219.27	15	409.4	27	242.30
4	152.92	16	452.85	28	411.30
5	122.09	17	357.40	29	191.48
6	49.88	18	460.50	30	312.80
7	127.00	19	309.20	31	383.60
8	299.77	20	200.50	32	177.10
9	304.47	21	160.09	33	67.36
10	78.00	22	45.56	34	82.29
11	56.76	23	455.80	35	293.16
12	90.83	24	225.60	36	48.40

يتضح من الجدول السابق أن قيمة كا^2 لجميع المفردات اكبر من القيم الجدولية عند مستوي (0.01) ، ويدل هذا علي المفردات داله إحصائية عند مستوي (0.01) .

ومن خلال هذه الخطوة ، تم التأكد من أن أسئلة المقياس قادرة علي التمييز ، بمعني الي أن كل سؤال يؤدي إلي اختلاف في الاستجابات ، ولم يتم حذف شئ لان جميع أسئلة المقياس لها القدرة علي التميز .

رابعاً تقنين المقياس :

بعد الإعداد لمفردات المقياس كان علي الباحث أن يتأكد من مدي صدق المقياس وثباته قبل تطبيقه علي عينة البحث الأصلية .

1) صدق المقياس :

الاختبار الصادق يقيس ما وضع لقياسه . وتتعدد أساليب حساب وتقدير الصدق ، فنحصل في بعض الحالات علي معامل كمي للصدق بالتعبيرات الرقمية ، أو نحصل في حالات أخري علي تقدير كيفي للصدق بالتعبيرات اللفظية .

أ- التقدير الكيفي للصدق :

1- الصدق الظاهري :

يدل الصدق الظاهري علي المظهر العام للاختبار كوسيلة من وسائل القياس النفسي ، ويبدو الصدق الظاهري في مدي مناسبة الاختبار للمفحوصين ، وفي وضح تعليماته وصحة ترتيبها للخطوات الأساسية التي يتبعها المفحوص في فهمه للأسئلة وإجابته عنها ، وفي نوع الأسئلة ومدي صلاحيتها لإثارة الاستجابات المناسبة من المفحوصين .

وفي سبيل تحقيق الصدق الظاهري اعتمد الباحث علي آراء مجموعة من المحكمين المتخصصين في علم النفس والصحة النفسية (نسبة اتفاق 80 ٪) . هذا بجانب ما أسفر عنه التجريب المبدئي من وضوح الأسئلة وقدرتها علي التمييز (بحساب كا^2) .

2- صدق المحتوي (المضمون) :

يهدف صدق المحتوي أو الصدق المنطقي إلي الحكم علي مدي تمثيل الاختبار للميدان الذي يقيسه .

ويعتمد بناء الاختبارات الحديثة علي هذا النوع من الصدق في صياغة وإعداد الاختبارات المختلفة ، فيبدءون بتحليل المجال أو الميدان الاختباري أو الناحية التي يراد قياسها تحليلاً يكشف عن عناصرها المختلفة وأقسامها الرئيسية ثم بفصل كل قسم إلي أجزائه المختلفة وتقدر النسب المئوية لأجزاء كل قسم من هذه الأقسام وبذلك تصبح عملية اختيار الأسئلة عملية صحيحة شاملة .

وفي سبيل تحقيق هذا النوع من الصدق فقد استعرض الباحث الدراسات السابقة في موضوع الحكمة الاختبارية ، بالإضافة إلي الاستفتاء المفتوح علي طلاب الصف الأول الثانوي العام . واستنتج الاستراتيجيات المستخدمة في قياس الحكمة الاختبارية ، ثم عرفها الباحث وعرضها هي والأسئلة الموضوعية عليها علي المحكمين ، وبذلك نضمن أن المقياس شامل للمجال الذي يقيسه .

ب – التقدير الكمي للصدق (الصدق الإحصائي) :

صدق المفردات :

وهو نوع من أنواع صدق التكوين الفرضي يتمثل في حساب صدق المفردات التي يتألف منها الاختبار ، وتعتبر الدرجة الكلية للاختبار هي نفسها محك للصدق .

ولتحقيق صدق التكوين الفرضي لمقياس الحكمة الاختبارية . فقد استخدم الباحث معامل الارتباط الثنائي .

ويعتمد معامل الارتباط الثنائي في حسابه علي الوصول إلي المتوسط الحسابي لكل من المتغيرين النوعين وعلي الانحراف المعياري للتكرارات الكلية .

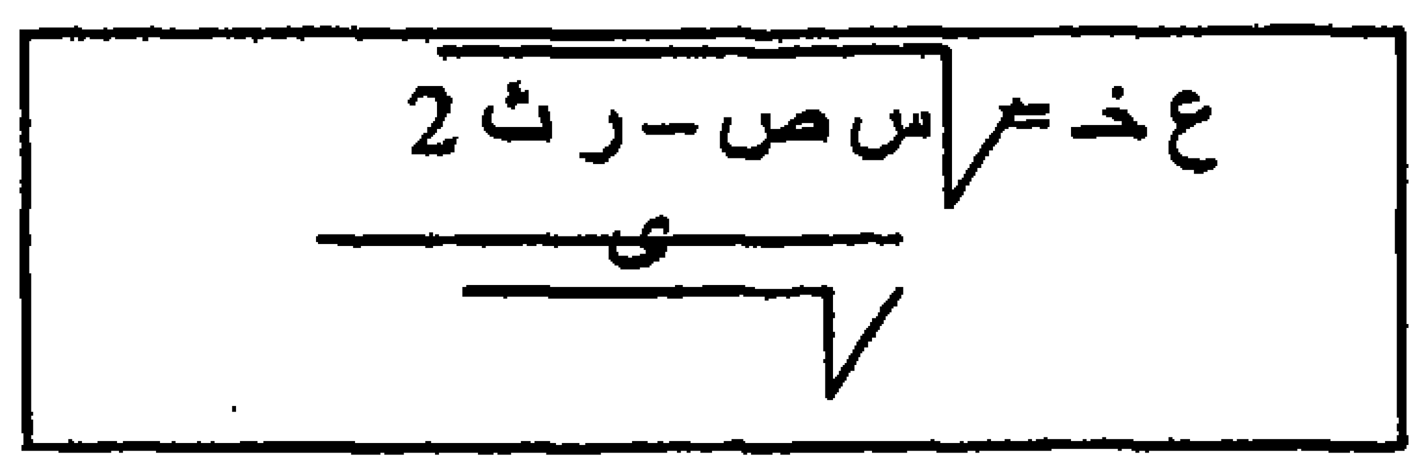
وقانون

$$ر = \frac{م_1 - م_2}{ع} \times \frac{أ . ب}{ص}$$

حيث يشير م1 إلي متوسط الصواب ، م2 ، أ إلي نسبة الصواب ب إلي نسبة الخطأ ، ص إلي الارتفاع الاعتدالي المقابل لنسبة الصواب .

حساب الدلالة الإحصائية لمعامل الارتباط الثنائي : وفي سبيل ذلك اتبعت طريقة جيلفورد التي تتلخص في الأتي :

1. حساب الخطأ المعياري لمعامل الارتباط الثنائي : ع خـ من المعادلة التالية :



حيث ن عدد أفراد العينة .

2- ضرب ع خـ × 1.96 (مستوي الدلالة عند 0.05) او ع خـ × 2.58 (مستوي الدلالة عند 0.01) فإذا كان :

(أ) 1.96 × ع خـ > رث كان رث غير دال إحصائياً ويجب حذف السؤال .

(ب) 1.96×ع خـ < رث < 2.58 × ع خـ فان رث يكون دالا إحصائياً عند مستوي (0.05) ، وبالتالي لا يحذف السؤال .

(ج) رث > 2.58 × ع خـ فان رث يكون دالا إحصائياً عند مستوي (0.01) ، ولا يحذف السؤال .

ويوضح الجدول رقم (5) معاملات الارتباط الثنائي ، ومستوي الدلالة

جدول رقم (5)

صدق مفردات مقياس الحكمة الاختبارية باستخدام معامل الارتباط الثنائي .

السؤال	الصواب			الخطأ			الارتفاع الاعتدالي	رث	ع غ	الدلالة
	عدد	متوسط	نسبة	عدد	متوسط	نسبة				
1	163	41.85	0.99	2	38.50	0.01	0.03	0.35	0.04	0.01
2	156	41.77	0.95	9	42.33	0.05	0.11	0.07 -	0.05	-
3	123	41.93	0.75	42	41.45	0.25	0.32	0.08	0.06	-
4	162	41.88	0.98	3	37.33	0.06	0.05	0.48	0.03	0.01
5	100	42.61	0.61	65	40.57	0.34	0.38	0.35	0.05	0.01
6	51	42.25	0.31	114	41.61	0.69	0.35	0.11	0.06	-
7	74	44.38	0.28	118	40.78	0.72	0.34	0.59	0.03	0.01
8	135	42.42	0.82	30	39.03	0.18	0.16	0.52	0.04	0.01
9	16	40.75	0.10	149	41.91	0.90	0.17	0.16 -	0.05	-
10	164	42.38	0.99	1	38.00	001.	0.02	0.37	0.04	0.01
11	93	42.43	0.56	72	41.00	.0.44	0.39	0.24	0.06	0.01
12	83	42.75	0.50	82	40.85	0.50	0.40	0.33	0.05	-
13	115	42.44	0.70	50	40.34	0.30	0.35	0.35	0.05	0.01
14	138	42.15	0.84	27	40.07	0.16.	0.25	0.31	0.05	0.01
15	154	41.99	0.93	11	39.18	0.07	0.13	0.37	0.04	0.01
16	161	41.90	0.98	4	37.75	0.02	0.06	0.48	0.03	0.01
17	147	41.84	0.89	18	41.56	0.11	0.19	0.04	0.06	0.01
18	161	41.95	0.98	4	36.00	0.02	0.06	0.68	0.01	0.01
19	158	42.06	0.96	7	36.00	0.04	0.09	0.75	0.01	0.01
20	157	41.88	0.95	8	40.38	0.05	0.010	0.19	0.05	0.01
21	98	42.17	0.59	67	41.57	0.41	0.39	0.10	0.06	-
22	128	46.08	0.78	37	40.84	0.22	0.30	0.20	0.05	0.01
23	159	41.35	0.96	6	40.67	0.04	0.08	0.08	0.05	-
24	147	42.20	0.89	18	38.44	0.11	0.19	0.54	0.03	0.01
25	84	42.71	0.51	81	40.86	0.49	0.40	0.32	0.05	0.01
26	145	42.05	0.88	20	40.05	0.12	0.20	0.29	0.05	0.01
27	164	41.81	0.99	1	41.00	0.01	0.17	0.08	0.05	-
28	165	41.81	1.00	-	-	-	-	0	-	-
29	117	42.01	0.71	48	41.29	0.29	0.34	0.11	0.06.	-
30	142	42.28	0.86	23	38.87	0.14	0.22	0.51	0.04	0.01
31	156	42.00	0.95	9	38.44	0.05	0.11	0.45	0.04	0.01
32	122	24.63	0.68	53	40.08	0.32	0.36	0.43	0.05	0.01
33	82	42.67	0.50	83	40.95	0.50	0.40	0.30	0.05	0.01
34	69	42.49	0.42	96	41.31	0.58	0.39	0.20	0.06.	0.01
35	158	42.01	0.96	7	37.14	0.04	0.09	0.60	0.02	0.01
36	72	42.11	0.44	93	41.57	0.56	0.39	0.09	0.06	-

وبهذه الطريقة حسب معامل الارتباط الثنائي لكل سؤال من المقياس بالدرجة الكلية سؤال من المقياس بالدرجة الكلية للمقياس وقد كشفت الارتباطات عن احد عشر سؤالاً أرقامها (2 ، 3 ، 6 ، 9 ، 17 ، 21 ، 23 ، 27 ، 28 ، 36)

لم تظهر ارتباطا عاليا مع الدرجة الكلية للمقياس باعتبارها لا تتمتع بقدر كاف من الصدق والصلاحية للقياس . وبذلك يتضح أن باقي الأسئلة (عددها 25 سؤالاً) علي درجة كبيرة من الاتساق والتجانس وصالحة لقياس ما وضع المقياس من أجله .

الصدق العاملي لمقياس الحكمة الاختبارية :

تم إجراء الارتباطات بين أسئلة المقياس بعضها والبعض الأخر بينها وبين الدرجة الكلية للمقياس . وقد بلغ عدد الارتباطات في مصفوفة الارتباط الأولي (666) معامل ارتباط . ولقد اجري التحليل العاملي لهذه الأسئلة والتي عولجت إحصائياً باعتيارها متغيرات بطريقة المكونات الأساسية لهوتلنج علي معاملات الارتباط بين هذه الأسئلة والتي عددها ستة وثلاثون سؤالاً ، وحسبت تشبعات الأسئلة علي العوامل قبل التدوير وبعد التدوير . وسيتم عرض ذلك تفصيلياً في الصفحات القادمة الخاصة بالتحليل العاملي لمقياس الحكمة الاختبارية.

2. ثبات مقياس الحكمة الاختبارية :

أ- باستخدام معامل " الفا " (معادلة كرونباك) :

يستخدم معامل الفا في حساب ثبات الاختبارات والمقاييس إذا كانت درجات المفردات تخرج عن نطاق الواحد الصحيح أو الصفر، كما أنه يصلح أيضا

إذا كانت درجات المفردات واحدا أو صفر وتمثل القيمة الناتجة من حساب ثبات الاختبار بمعامل الفا الحد الأدنى لثبات الاختبار .

ولقد استخدمت المقاييس الأجنبية السابقة في الحكمة الاختبارية معادلة كيودر ريتشاردسون (K-R20) ، ومعامل الف في حساب معامل الثبات ، وأن كانت معظمها أظهرت قيما منخفضة لمعامل الثبات .

ولما كانت المقاييس الأجنبية السابقة تدور حول إستراتيجية واحدة من استراتيجيات الحكمة الاختبارية أو علي الأكثر إستراتيجيتين ، وفيها الإجابات تأخذ صفرا أو واحدا . فمن الممكن استخدام معادلة كيودر يتشاردسون ، لأنها تستخدم عندما تأخذ الإجابات صفرا وواحدا . ونظرا لأن إجابات أسئلة مقياس الدراسة الحالية تأخذ أوزانا (صفر ، واحد) ، (صفر ، 1 ، 2) ،(صفر ، 1 ، 2 ، 3) ، فقد استخدم معامل ألفا (معادلة كرونباك) ، وهذه طريقة عامة لقياس معامل الثبات . فعندما تأخذ (ن) أكبر قيمة ممكنة ، وهي عدد وحدات الاختبار ، وبذلك نجعل الجزء أصغر ما يمكن ، وبهذا نحصل علي معادلة (K-R20) ، وإذا جعلنا (ن) تأخذ أصغر قيمة ممكنة فأننا نحصل علي معادلة جتمان .

وفي سبيل ذلك فقد استخدم معامل الفا (معادلة كرونباك) في حساب معامل الثبات للمقياس وذلك قبل الحذف للأسئلة التي لم تكن دالة بالنسبة لصدق المفردات . وتأخذ المعادلة الصيغة :

$$\text{ر}_{\text{اا}} = \frac{\text{ن}}{\text{ن}-1}\left[\frac{1-\text{مجـ ع}^2\text{ك}}{\text{ع}^2}\right]$$

حيث أن ر أأ معامل الثبات ، ن عدد الأسئلة (الأجزاء) ، ع 2 ك تباين الجزء ك من أجزاء الاختبار ، ع 2 تباين الاختبار كله ، مجـ ع2ك مجموع تباينات الأجزاء كلها .

كما حسب معامل الثبات بمعامل الفا بعد حذف أحد عشر سؤالا (التي ليست علي قدر كاف من الصدق) . وكانت قيم معامل الثبات كما يلي :

معامل الثبات قبل حذف أحد عشر سؤالا = 0.35

معامل الثبات قبل حذف أحد عشر سؤالا = 0.37

ولما كانت قيم معامل ألفا منخفضة ، وهذا ما أشارت إليه بعض المقاييس في الدراسات الأجنبية السابقة ، فقد حسب الباحث معامل الفا لمفردات كل إستراتيجية من استراتيجيات الحكمة الاختبارية الست في المقياس ، ويوضح الجدول التالي قيم معامل ألفا لمفردات كل إستراتيجية من استراتيجيات الحكمة الاختبارية .

جدول رقم (6)

معاملات ألفا لاستراتيجيات الحكمة الاختبارية في المقياس

الإستراتيجية	عدد الأسئلة	معامل الفا
استخدام دليل أو رمز	14	0.21
الاستدلال الاستنباطي	9	0.11
استخدام الوقت	4	0.05
تجنب الخطأ	3	0.14
التخمين	3	0.04
مراعاة القصد أو الهدف	3	0.06

ولقد قام الباحث بتطبيق المقياس علي فصلين ، أحدهما للبنين والآخر للبنات ويوضح ذلك الجدول رقم (7) ، ثم إعادة تطبيق مقياس الحكمة الاختبارية بعد أسبوعين علي نفس الفصلين بغرض حساب معامل الثبات بطريقة الإعادة ، حيث أن قيمة معامل الثبات بطريقة الإعادة في المقاييس السابقة الأجنبية كانت أعلي من قيمته المحسوبة بواسطة معامل الفا أو كيودر يتشاردسون (20) ، وفي سبيل ذلك فقد حسب معامل الارتباط بين درجات التطبيقين الأول والثاني ، وذلك قبل الحذف . ويوضح الجدولين التاليين درجات التطبيق الأول والثاني ، وذلك قبل الحذف . ويوضح الجدولين التاليين درجات العينة التي حسب عليها معامل الثبات قبل الحذف . وبعد الحذف .

جدول رقم (7)

درجات العينة التي حسب عليها معامل ثبات مقياس الحكمة الاختبارية بطريقة الإعادة وقبل حذف الأسئلة التي صدقها غير دال إحصائياً

م	التطبيق الأول	التطبيق الثاني	م	التطبيق الأول	التطبيق الثاني	م	التطبيق الأول	التطبيق الثاني
1	41	42	24	39	40	47	49	44
2	43	41	25	32	43	48	41	37
3	38	32	26	41	43	49	45	45
4	36	41	27	35	32	50	44	47
5	35	35	28	36	39	51	44	42
6	38	40	29	36	31	52	46	41
7	43	39	30	35	32	53	45	43
8	36	43	31	28	23	54	45	42
9	43	40	32	36	43	55	41	43
10	46	36	33	42	42	56	40	37
11	41	40	34	39	32	57	41	46
12	38	41	35	46	42	58	48	48
13	40	42	36	40	44	59	44	46

140 14	39	39	37	45	48	60	41	40
15	38	35	38	49	45	61	47	47
16	36	45	39	44	46	62	42	46
17	40	36	40	40	34	63	40	42
18	42	30	41	43	45	64	41	42
19	43	42	42	47	44	65	42	43
20	43	35	43	45	48	66	42	44
21	31	45	44	40	45	67	46	46
22	40	43	45	36	37	68	45	48
23	40	41	46	41	46	69	43	45

جدول رقم (8)

درجات العينة التي حسب عليها معامل ثبات مقياس الحكمة الاختبارية بطريقة الإعادة وقبل حذف الأسئلة التي صدقها غير دال إحصائياً

م	التطبيق الأول	التطبيق الثاني	م	التطبيق الأول	التطبيق الثاني	م	التطبيق الأول	التطبيق الثاني
1	28	31	24	30	40	47	37	32
2	32	31	25	23	43	48	29	28
3	28	26	26	30	43	49	32	33
4	26	30	27	26	32	50	32	33
5	23	25	28	26	39	51	33	32
6	30	28	29	23	31	52	33	31
7	29	28	30	26	32	53	33	33
8	25	31	31	19	23	54	32	30
9	32	28	32	26	43	55	31	32
10	33	27	33	30	42	56	30	26
11	32	32	34	29	32	57	30	34
12	29	30	35	35	42	58	35	26
13	31	31	36	33	44	59	32	32

14	30	28	37	34	48	60	29	28
15	27	2	38	36	45	61	37	36
16	29	32	39	33	46	62	33	35
17	29	24	40	26	34	63	27	28
18	32	23	41	31	45	64	32	32
19	34	33	42	36	44	65	30	31
20	30	27	43	33	48	66	35	24
21	21	32	44	30	45	67	33	22
22	29	31	45	29	37	68	33	35
23	31	32	46	30	31	69	30	33

وكانت قيم معامل الثبات للمقياس قبل الحذف = 0.56 ، وبعد الحذف = 0.68 ، وبذلك تكون قيمة معامل الثبات بطريقة الإعادة للمقياس وبعد حذف الأسئلة التي كان صدقها غير دال إحصائيا معقولة .

جدول رقم (9)

العينة التي حسب عليها معامل الثبات بطريقة الإعادة لمقياس الحكمة الاختبارية

المدرسة	الفصل	العدد	الإجمالي
احمد عرابي الثانوية بنين	2/1	34	69
الزقازيق الثانوية بنات	1/1	35	

ثبات مفردات مقياس الحكمة الاختبارية :

يعتمد ثبات الاختبار اعتمادا مباشرا علي ثبات مفرداته وتتخلص أهم الطرق الإحصائية لحساب ثبات المفردات في طريقة إعادة الاختبار وطريقة الاحتمالي المتوالي .

وقد استخدم الباحث طريقة الاحتمال المنوالي لحساب ثبات المفردات ، وذلك لأن هذه الطرقة تصلح لحساب ثبات المفردات التي تعتمد إجابتها علي اختيار إجابة واحدة من عدة إجابات محتملة ، كما تصلح أيضا لحساب ثبات الأسئلة التي تقوم فكرتها علي الاحتمال الاختياري .

وتتخلص معادلة الثبات في الصورة التالية :

$$\text{معامل الثبات} = \frac{ن}{ن-1}\left(ل - \frac{1}{ن}\right)$$

حيث يدل الرمز (ن) علي عدد الاحتمالات الاختيارية للسؤال ، ويدل الرمز (ل) علي الاحتمال المنوالي اي علي أكبر تكرار نسبي لأي احتمال اختياري من الاحتمالات التي يحتوي عليها السؤال .

وقد استخدمت الطريقة الآتية في الكشف عن الدلالة الإحصائية للفرق بين النسبتين (الاحتمال المنوالي والذي يساوي أكبر تكرار مقسوما علي مجموع التكرارات ، والقيمة الثابتة وتساوي (1) . وذلك كما يلي :

1. تحسب قيمتي $ل_1$ ، $ل_2$ كما يلي :

$ل_1 = 2$ جا$^{-1}$ ب ، $ل_2 = 2$ جا$^{-1}$ جـ

حيث (ب) النسبة الأولي وتساوي أكبر تكرار مقسوماً علي مجموع التكرارات ، (جـ) النسبة الثانية وتساوي:

وهي في مقياس الدراسة $= \frac{1}{ن} = 0.25$

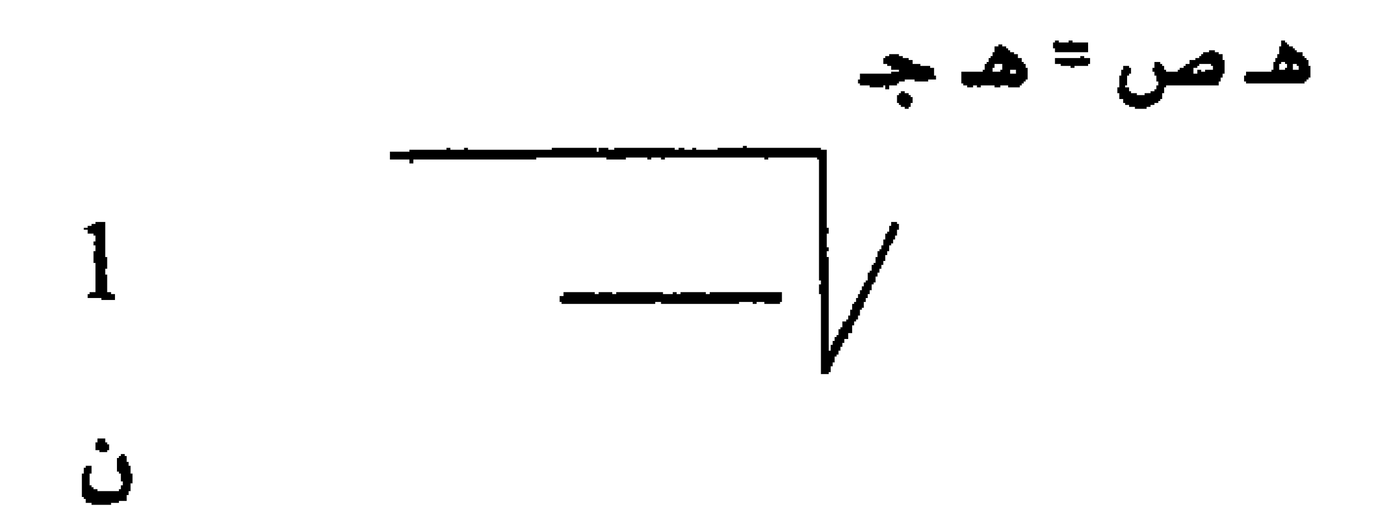
لأن عدد الاحتمالات الاختيارية في المقياس $(4)^{\text{ن}}$

تحسب قيمة (هـ س) كما يلي : هـ س = $\text{ل}_1 - \text{ل}_2$

2. يكشف عن قيمة (هـ جـ) بدلالة عدد الأفراد (ن) للنسبة الأولي .

3. تحسب قيمة (هـ ص) كما يلي :

هـ ص = هـ جـ



$$\sqrt{\frac{1}{\text{ن}}}$$

4. نقارن قيمة (هـ س) بقيمة (هـ ص) فإذا كانت : هـ س > هـ ص كان الفرق بين النسبتين دالا إحصائياً ، وبالتالي يكون معامل ثبات المفردات دالا إحصائياً .

ويوضح الجدول رقم (10) حساب معاملات ثبات مفردات المقياس ، وقيم هـ س ، هـ ص ، ومستوي الدلالة الإحصائية للفرق بين النسبتين ، ومعاملات الثبات .

جدول رقم (10)

معاملات ثبات المفردات في مقياس الحكمة الاختبارية ومستوي دلالتها الإحصائية

رقم السؤال	معامل الثبات	هـ س	هـ ص	رقم السؤال	معامل الثبات	هـ س	هـ ص
1	0.63	0.98	0.20	19	0.79	1.27	0.20
2	0.93	1.62	0.20	20	0.60	0.94	0.20
3	0.65	1.02	0.20	21	0.47	0.73	0.20
4	0.96	1.75	0.20	22	0.96	0.26	0.20
5	0.49	0.77	0.20	23	0.96	1.75	0.20
6	0.31	0.34	0.20	24	0.67	1.05	0.20
7	0.47	0.73	0.20	25	0.31	0.48	0.20
8	0.78	1.26	0.20	26	0.85	1.72	0.20
9	0.77	1.25	0.20	27	0.68	1.07	0.20
10	0.83	1.16	0.20	28	0.91	1.56	0.20
11	0.39	0.61	0.20	29	0.61	0.96	0.20
12	0.34	0.52	0.20	30	0.80	1.30	0.20
13	0.39	0.61	0.20	31	0.88	1.49	0.20
14	0.81	1.33	0.20	32	0.59	0.91	0.20
15	0.91	1.56	0.20	33	0.32	0.50	0.20
16	0.96	1.75	0.20	34	0.24	0.38	0.20
17	0.84	1.39	0.20	35	0.77	1.25	0.20
18	0.97	1.78	0.20	36	0.24	0.38	0.20

علما بأن مستوي دلالة الفرق بين النسبتين لجميع المفردات (0.01) ، ومستوي دلالة معامل الثبات لجميع المفردات ايضا (0.01)

خامسا : التحليل العاملي لمقياس الحكمة الاختبارية :

بعد تطبيق مقياس الحكمة الاختبارية علي العينة البالغ عددها (165) طالبا وطالبة دونت درجات الطلاب في الكشوف الخاصة بذلك واستخرجت

الارتباطات بين أسئلة المقياس بعضها والبعض الآخر ، وبينها وبين الدرجة الكلية للمقياس ، ثم تم حساب التحليل العاملي بطريقة المكونات الأساسية Principle Components Factor Analysis بين هذه الارتباطات وقد سارت خطة التحليل العاملي في اتجاهين :

أولهما : التحليل العاملي لمتغيرات المقياس والدرجة الكلية .

ثانيهما : التحليل العاملي لمتغيرات (مفردات) المقياس .

أولا : التحليل العاملي الأول لمتغيرات المقياس والدرجة الكلية :

- الارتباطات بين الدرجة الكلية لمقياس الحكمة الاختبارية ومفردات المقياس :

1. أن الدرجة الكلية للمقياس ترتبط ارتباطا موجبا بكل سؤال من أسئلة المقياس ماعدا السؤالين الثاني والتاسع . فنجد أن ارتباط الدرجة الكلية بكل منهما سالب ، وهذا يتفق مع معامل الارتباط الثنائي في حذفه للسؤالين الثاني والتاسع .

2. أن الارتباطات الموجبة بين الدرجة الكلية للمقياس والأسئلة معظمها دالة ماعدا الأسئلة (1 ، 3 ، 6 ، 7 ، 17 ، 20 ، 21 ، 23 ، 29 ، 36) ، فارتباطاتها موجبة لكنها ضعيفة ، وهذا متفق تقريباً مع معامل الارتباط الثنائي في استبعاده للأسئلة (3 ، 6 ، 17 ، 21 ، 23 ، 29 ، 36) . وبذلك نجد أن الأسئلة استبعدت بواسطة معامل الارتباط الثنائي عند حساب صدق المفردات للمقياس هي نفسها تقريباً التي ظهرت ارتباطاتها بالدرجة الكلية للمقياس اما سالبة ضعيفة (غير دالة) أو موجبة ضعيفة (غير دالة) ، وأما بقية الأسئلة فتتمتع بارتباطات داله إحصائية بالدرجة الكلية للمقياس.

ثانياً : الارتباطات بين أسئلة المقياس بعضها البعض :

3) أن المفردة الأولي (الإجابة بسرعة مع الإتقان) ترتبط ارتباطا موجبا دالا مع المفردة العاشرة (توقع التضليل والتعقيد) ، وترتبط ارتباطا سالبا دالا مع المفردة الحالية عشرة (الاختيارات غير المقبولة) .

4) ترتبط المفردة الثانية (الاختيارات المتشابهة) ارتباطا سالبا دالا بالمفردتين الحادية والعشرين (التحديات الخاصة) والثالثة والثلاثين (الاستفادة من معلومات محتوي مشابه) .

5) ترتبط المفردة الثالثة (الاختيارات المتفرعة) ارتباطا موجبا دالا بالمفردتين التاسعة عشرة (استخدام التخمين) والثالثة والعشرين (التحديات الخاصة) ، وارتباطا دالا بالمفردات الثانية والعشرين (إجابة المفردات وفقا لهدف الاختبار ومصمم الاختبار) والمفردة الثلاثين (التحديدات الخاصة) والمفردة الخامسة والثلاثين (مراعاة القصد والهدف) .

6- ترتبط المفردة الرابعة (قراءة التعليمات ارتباطا موجبا دالا بالمفردات ، التاسعة عشرة (استخدام التخمين) والرابعة والعشرين (إرجاء أو تخمين المفردات التي لا يستطيع المفحوص الإجابة عليها بسرعة) ، والثلاثين (التحديات الخاصة) ، كما ترتبط ارتباطا سالبا دالا بالمفردة التاسعة (الاختيار الصحيح يكون متناسقا لغويا مع الجذع) .

7- ترتبط المفردة الخامسة (الاختيارات المتشابهة) ارتباطا موجبا دالا بالمفردات ، الثامنة (الاختيارات المتشابهة) والثانية والثلاثين (الاختيارات المتداخلة) ، كما ترتبط ارتباطا سالبا دالا بالمفردة الثانية والعشرين (إجابة المفردات وفقا لقصد مصمم الاختبار والهدف منه) .

8- ترتبط المفردة السادسة (الأخذ في الاعتبار مادة الموضوع) ارتباطا موجبا دالا بالمفردات العاشرة (توقع التضليل أو التعقيد) ، والخامسة والعشرين (الاستفادة من معلومات محتوي مشابه) ، والثانية والثلاثين(الاختيارات المتداخلة) ، كما ترتبط ارتباطا سالبا دالا بالمفردات الرابعة عشرة (الاختيارات غير المقبولة) ، السابعة عشرة (الاختيارات غير المقبولة) ، الثامنة عشرة (التحديات الخاصة) ، السابعة والعشرين (مراجعة جميع الإجابات) ، والتاسعة والعشرين(استخدام التخمين) ، السادسة والثلاثين (الاختيارات المتداخلة).

9- ترتبط المفردة السابعة (استخدام التخمين ارتباطا موجبا دالا بالمفردة الثانية عشرة (الاختيار الصحيح يكون متناسقا لغويا مع الجذع) ، كما ترتبط ارتباطا سالبا بالمفردات الرابعة عشرة والثانية والعشرين .

10- ترتبط المفردة الثامنة (الاختيارات المتشابهة) ارتباطا موجبا دالا مع المفردات التاسعة عشرة ، الثلاثين ، الحادية والثلاثين (الإفادة من الزمن المتبقي في مراجعة الإجابات) ، الثالثة والثلاثين (الاستفادة من معلومات محتوي مشابه) وترتبط ارتباطا سالبا دالا بالمفردة التاسعة .

11- ترتبط المفردة التاسعة (الاختيار الصحيح يكون متناسقا لغويا مع الجذع) ارتباطا سالبا دالا مع المفردات الربعة عشرة ، الثامنة عشرة ، التاسعة عشرة ، الثانية والعشرين .

12- ترتبط المفردة العاشرة (توقع التضليل أو التعقيد) ، ارتباطا موجبا دالا مع المفردة الخامسة والعشرين كما ترتبط ارتباطا سالبا دالا بالمفردات الثامنة عشرة ، والتاسعة والعشرين ، والثلاثين .

13- ترتبط المفردة الحادية عشرة (الاختيارات غير المقبولة ارتباطا موجبا دالا بالمفردة الحادية والعشرين .

14- ترتبط المفردة الثانية عشرة (الاختيار الصحيح يكون متناسقا لغويا مع الجذع) ارتباطا موجبا دالا بالمفردات الثالثة عشرة (وضع قائمة أو جدول لما يحدث أثناء الاختيار) ، وبالمفردة الثالثة والثلاثين (الاستفادة من معلومات محتوي مشابه).

15- ترتبط المفردة الثالثة عشرة ارتباطا سالبا دالا بالمفردة الثانية والعشرين .

16- ترتبط المفردة الرابعة عشرة (الاختيارات غير المقبولة) ارتباطا موجبا دالا بالمفردات السابعة عشرة ، والثانية والعشرين ، السابعة والعشرين ، الثامنة والعشرين ، التاسعة والعشرين ، الخامسة والثلاثين ، كما ترتبط ارتباطا سالبا دالا بالمفردة الرابعة والثلاثين .

17- ترتيط المفردة الخامسة عشرة (الاختيار الصحيح يكون متناسقا لغويا مع الجذع) ارتباطا موجبا دالا بالمفردات (25 ، 30 ، 31 ، 32 ، 33)

18- لا ترتبط المفردة السادسة عشرة (قراءة المفردات بعناية) بأي مفردة من مفردات المقياس بينما ترتبط ارتباطا موجبا دالا بالدرجة الكلية للمقياس .

19- ترتبط المفردة السابعة عشرة (الاختيارات غير المقبولة) ارتباطا موجبا دالا بالمفردتين (18 ، 29) وسالبا بالمفردة (25) .

20- ترتبط المفردة الثامنة عشرة (التحديات الخاصة) ارتباطا موجبا دالا بالمفردات (19 ، 23 ، 26 ، 28 ، 29 ، 30 ، 33) .

21- ترتبط المفردة التاسعة عشرة (استخدام التخمين) ارتباطا موجبا دالا بالمفردات (23 ، 24 ، 27 ، 30 ، 42)

22- ترتبط المفردة (20) (عدم اختيار المفردتين المضادتين لبعضهما) ارتباطا سالبا دالا بالمفردة (24) (إرجاء وتخمين المفردات التي لا يستطيع المفحوص الإجابة عليها بسرعة) .

23- ترتبط المفردة (21) (التحديات الخاصة) ارتباطا موجبا دالا بالمفردة (28) ، وسالبا بالمفردة (24) .

24- ترتبط المفردة (22) (إجابة المفردات وفقا لقصد مصمم وهدف الاختبار) ارتباطا موجبا دالا بالمفردة (30) (التحديات الخاصة) .

25- ترتبط المفردة (25) (الاستفادة من معلومات محتوي مشابه) ارتباطا موجبا دالا بالمفردة (33) ، وارتباطا سالبا دالا بالمفردة (27) (مراجعة جميع الإجابات) .

26- ترتبط المفردة (26) (مراجعة جميع الإجابات) ارتباطا موجبا دالا بالمفردة (28) (الترابط بين البديل الصحيح والجذع) .

27- ترتبط المفردة (28) (الترابط بين البديل الصحيح والجذع) ارتباطا موجبا دالا بالمفردة (35) (مراعاة القصد أو الهدف) .

28- ترتبط المفردة (30) (التحديات الخاصة) ارتباطا موجبا دالا بالمفردة (35) .

29- ترتبط المفردة (33) (الاستفادة من معلومات محتوي مشابه) ارتباطا موجبا دالا بالمفردة (35)

(مراعاة القصد أو الهدف) .

ولقد تم إجراء تحليل عاملي من الدرجة الأولي بطريقة المكونات الأساسية لهونج وذلك باستخدام قاعدة كايزر Kaiser Rule التي مؤداها اختيار

العوامل ذات القيم الكامنة (الجذور الكامنة) الأكبر من الواحد الصحيح أو تساويه ووصل عدد العوامل المستخلصة إلي (15) عاملا .

وهذا الحل الأولي فسر (66.40٪) من التباين الكلي للستة والثلاثين متغيرا (سؤالا).

وعن تشبعات مفردات مقياس الحكمة الاختبارية علي العوامل الخمسة عشر قبل التدوير، يتضح ما يلي:-

1. المفردات (4 ، 8 ، 9 ، 14 ، 18 ،19، 24 ، 27 ، 30) ، والدرجة الكلية تشبعاتها بالعامل الأول تزيد عن 30٪ ،وهذه القيمة دالة إحصائيا وعمليا .

2. كما أن المفردات (5 ،6 ، 7 ،12، 9، 13، 14، 25، 22، 17، 15، 29، 33 ، 34) والدرجة الكلية تشبعاتها بالعامل الثاني دالة إحصائيا وعمليا (30 ٪فأكثر).

3. ووجد أن المفردات (2 ،3، 6 ، 19، 21، 23 ، 28، 33) تشبعاتها بالعامل الثالث دالة إحصائيا وعمليا (30 ٪فأكثر) .

4. وتشبعات المفردات (1 ، 10 ، 21 ،36) بالعامل الرابع دالة إحصائيا (30 ٪ فأكثر)

5. وتشبعات المفردات (2 ،4، 16، 18، 30) بالعامل الخامس دالة إحصائيا (30 ٪ فأكثر)

6. وتشبعات المفردات (10 ، 11، 14 ، 21) بالعامل السادس دالة إحصائيا (30 ٪ فأكثر)

7. وتشبعات المفردات (15 ، 18 ، 31 ، 32) بالعامل السابع دالة إحصائيا (30 ٪ فأكثر)

8. وتشبعات المفردات (8، 25، 31 ، 33) بالعامل الثامن دالة إحصائيا (30 ٪ فاكثر)

9. وتشبع المفردة (34) بالعامل التاسع دالة إحصائيا (30 ٪ فاكثر)

10. وتشبعات المفردات (17 ، 18، 35) بالعامل العاشر دالة إحصائيا (30 ٪ فاكثر)

11. وتشبع المفردة (8) بالعامل الحادي عشر دالة إحصائيا (30 ٪ فاكثر)

12. وتشبعت المفردتين (5 ،7) بالعامل الثاني عشر دالة إحصائيا وعمليا (30 ٪ فاكثر)

13. وتشبعت المفردتين (8 ،36) بالعامل الثالث عشر دالة إحصائيا وعمليا (30 ٪ فاكثر)

14. وتشبعات المفردات علي العامل الرابع عشر غير دالة إحصائيا وعمليا (اقل من 30٪) ، اما المفردة السابعة فتشبعها بالعامل الخامس عشر دال إحصائيا وعمليا (30٪ فاكثر)

وللحصول علي حل نهائي قابل للتفسير ، فقد تم تدوير العوامل الخمسة عشر تدويرا متعامدا بطريقة الفاريماكس Rotion Varimax orthogonal ، وهذا التدوير مكن الباحث من افتراض انه يمكن إيجاد حل عاملي ذي معني من خلال ثمانية او تسعة عوامل فاجري التحليل العاملي مرة أخري بافتراض أن عدد العوامل هو ثمانية ، ثم بافتراض أن عدد العوامل تسعة ،وفي كل مرة تم تدوير المحاور بطريقة الفاريماكس . الحل الأخير كان لغرض التشخيص من حيث إمكانية الحصول علي عامل تاسع بقيمة كامنة اكبر من او تساوي واحدا وأن

يكون له معني ، ولم يتم الحصول علي هذا العامل ، واكتفي الباحث بحل عاملي ذي ثمانية عوامل ، هذه العوامل الثمانية حققت ثلاثة شروط أساسية هي :-

الأول : أن القيم الكامنة لكل عامل اكبر من أو تساوي واحدا (قاعدة كايزر) .

الثاني : أن معامل الارتباط بين المتغير والعامل (التشبع) أكبر من أو يساوي (0.3) لاعتبار أن هذا الحد من الارتباط يعني بالدلالة الإحصائية والعملية للتشبعات .

الثالث : أن هذه العوامل الثمانية أعطت تفسيرا مقبولا لا يحقق أبسط بناء عاملي .

ويتضح من مصفوفة العوامل بعد التدوير المتعامد ما يلي :

أن هذه العوامل طائفية :

العامل الأول :

وقد تشبع بهذا العامل (6) مفردات من المقياس بالإضافة إلي الدرجة الكلية للمقياس ، إذا نظرنا إلي مضمون هذه المفردات نلاحظ انها تشير إلي :تجنب الخطأ في صورة قراءة التعليمات واستخدام التخمين ومراعاة القصد والهدف والتعرف علي التحديدات الخاصة المرتبطة بالإجابة الصحيحة وغير الصحيحة وجعل الإجابة الصحيحة متناسقة لغويا مع الجذع ، ويمكن انطلق علي هذا العامل " تجنب الخطأ في مقابل جعل الإجابة الصحيحة متناسقة لغويا مع الجذع " ،وهو يستوعب (7.30) من التباين الكلي ،وهذا العامل قطبي .

العامل الثاني :-

وقد تشبع بهذا العامل ثلاث مفردات أرقامها (25 ، 33 ،35) وكلها تشبعات موجبة بالإضافة إلي الدرجة الكلية تشبعها بهذا العامل موجب ،وإذا نظرنا إلي مضمون هذه المفردات نلاحظ أنها تشير إلي الاستفادة من معلومات محتوي مشابه للاختبار في إجابة المفردات وينتمي هذا العنصر للاستدلال الاستنباطي ،ويمكن ان نطلق علي هذا العامل " استفادة المفحوص من معلومات محتوي مشابه للاختبار في الإجابة علي مفردات الاختبار المقدم له " وه يستوعب (5.90) من التباين الكلي .

العامل الثالث :-

وقد تشبع بهذا العامل خمس مفردات أرقامها (6 ، 14 ،27 ، 28 ، 29) بالإضافة إلي الدرجة الكلية للمقياس ، وكلها تشبعات موجبة ما عدا المفردة (6) تشبعها سالب ،وإذا نظرنا إلي مضمون هذه المفردات نلاحظ أنها تشير إلي تجنب الخطأ واستبعاد الاختيارات غير المقبولة (غير الصحيحة)والترابط بين البديل الصحيح والجذع واستخدام التخمين ،والأخذ في الاعتبار مادة الموضوع وصعوبة المفردات المجاورة عند تفسير وإجابة المفردة المقدمة للمفحوص ، ويمكن أن نطلق علي هذا العامل مراجعة جميع الإجابات في مقابل الأخذ في الاعتبار مادة الموضوع وصعوبة المفردات المجاورة عند تفسير وإجابة المفردة المقدمة للمفحوص . وهذا العامل قطبي ،ويستوعب (3.70) من التباين الكلي .

العامل الرابع :

وقد تشبع بهذا العامل مفردتان تشبعا موجبا أرقامهما (5 ،32) بالإضافة إلي الدرجة الكلية للمقياس ،وحين ننظر إلي مضمون كل منهما نلاحظ ان أحداهما تشير إلي استبعاد الاختيارات المتشابهة والأخرى تشير إلي الاختيارات

المتداخلة ، ويمكن أننطلق علي هذا العامل " استخدام الاختيارات المتشابهة والمتداخلة ، وهو يستوعب (3.4) من التباين الكلي .

العامل الخامس :

وقد تشبع بهذا العامل ثلاث مفردات أرقامها (14، 17، 29) ، وكلها تشبعات موجبة ،وقد تشبع بهذا العامل ثلاث مفردات أرقامها (14، 17، 29) ، وكلها تشبعات موجبة ،وإذا نظرنا إلي مضمون هذه المفردات نلاحظ أنها تشير إلي استبعاد الاختبارات التي من الواضح أنها غير صحيحة والاختيار من بين الاختيارات المتبقية ويطلق علي هذا العنصر استبعاد الاختيارات غير المقبولة ،وذلك يتبع الاستدلال الاستنباطي ، كما تشير إلي استخدام التخمين ، ويمكن أننطلق علي هذه العامل " استبعاد الاختيارات التي من الواضح أنها غير صحيحة والاختيار أو التخمين من بين الاختيارات المتبقية وهو يستوعب (2.9)من التباين الكلي.

العامل السادس :

وقد تشبع بهذا العامل ثلاث مفردات أرقامها (12 ،13 ، 24) بالإضافة إلي الدرجة الكلية وكلها تشبعات موجبة ،وإذا نظرنا إلي مضمون هذه المفردات نلاحظ أنها تشير إلي استخدام الوقت وجعل الاختيار الصحيح متناسقا لغويا مع الجذع ،ويمكن أننطلق علي هذا العامل " استخدام الوقت" وهو يستوعب (3.00) من التباين الكلي .

العامل السابع :-

وقد تشبع بهذا العامل ثلاث مفردات أرقامها (10 ،18 ، 23) ، منها التشبعات الموجبة (18 ،23) ،والمفردة (10) تشبعها سالب ،وإذا نظرنا إلي

مضمون هذه المفردات نلاحظ أنها تشير إلي التعرف علي التحديدات الخاصة المرتبطة بالاختيارات الصحيحة وغير الصحيحة ،وهذا العنصر ينتمي إلي إستراتيجية استخدام دليل أو رمز ، ومراعاة القصد أو الهدف للمفردة " وهو يستوعب (2.7) من التباين الكلي .

العامل الثامن :

وقد تشبع بهذا العامل ثلاث مفردات أرقامها (8 ،15، 31) وكلها تشبعات موجبة ،وإذا نظرنا إلي مضمون هذه المفردات والتشبعات الأعلى نلاحظ أنها تشير إلي الإفادة من الزمن المتبقي في مراجعة الإجابات ، وتناسق الاختيار الصحيح لغويا مع الجذع واستبعاد الاختيارات المتشابهة ،ويمكن أننطلق علي هذا العامل : " الإفادة من الزمن المتبقي في مراجعة الإجابات " .

وبذلك يقدم التحليل العاملي لمقياس الحكمة الاختبارية دليلا قويا علي مدي كفاءة مفردات هذا المقياس وصلاحيتها في قياس الجوانب التي وضعت من أجلها ، كما يتضح من

هذا التحليل أن العوامل التي تم استخلاصها تتفق مع الاستراتيجيات التي اعتمد عليها الباحث في بناء مقياس الحكمة الاختبارية فكل عامل من العوامل الثمانية الناتجة من التحليل العاملي بعد التدوير إما يمثل إستراتيجية من استراتيجيات الحكمة الاختبارية المحددة فى الإطار النظري أو يمثل أحد عناصر هذه الإستراتيجية ، كما نجد أن معظم المفردات علي العوامل ذات تشبعات دالة ومرتفعة.

ثانيا : التحليل العاملي لمتغيرات (مفردات)مقياس الحكمة الاختبارية :

ولقد أجري التحليل العاملي لمفردات المقياس الست والثلاثين بطريقة المكونات الأساسية المتعاقبة with interation principle component factoring وهذا النوع من التحليل العاملي (p A2) يستخدم فقط مربع معامل الارتباط المتعدد كتقدير أولي للاشتراكيات ، وأن هذه الاشتراكيات يعاد تقديرها مرات من خلال نظام حسابي معين Iteration procedure بهدف تحسين هذه التقديرات ، ومنطلق هذا النظام الحسابي المعين فهمه كما يلي :

1. تقدر الاشتراكيات تقديرا أوليا من خلال مربع معامل الارتباط المتعدد بين المتغير والمتغيرات الأخرى وذلك بعد تحديد عدد العوامل الممكنة .

2. بعد الخطوة الأولي السابقة يتم استخراج Extract نفس عدد العوامل السابقة من المصفوفة المتبقية Matrix Reduced ثم يعاد تقدير الاشتراكيات مرة أخري .

3. تكرر الخطوة (2) حتي نجد أن الفرق بين تقديرين متعاقبين للاشتراكيات يعد لا قيمة له بحيث يمكن إهماله.

ولقد استخدم الباحث طريقة المكونات الأساسية المتعاقبة في التحليل العاملي الثاني لمتغيرات (مفردات) مقياس الحكمة الاختبارية ،وتم استخراج عشر عاملا ، واستخدمت التقديرات في الحصول علي خمسة عشر عاملا وذلك باستخدام قيد إحصائي أخر وهو قاعدة كايزر والتي مؤداها اختيار العوامل ذات القيم الكامنة الأكبر أو تساوي واحد صحيحا .وهذه العوامل الخمسة عشر تمثل الحل الأولي وهذا الحل الأولي فسر (65.9) من التباين الكلي للست والثلاثين مفردة .

وكانت نتيجة الخطوة السابقة مصفوفة العوامل قبل التدوير والتشبعات الدالة إحصائيا وعمليا لمفردات مقياس الحكمة الاختبارية علي العوامل الخمسة عشر قبل التدوير حيث نجد أن :

1. تشبعات المفردات (6، 9، 10 ، 14 ، 15، 117، 18، 22، 25، 27 ، 29، 33) بالعامل الأول دالة إحصائيا وعمليا (30 ٪فاكثر) .

2. تشبع بالعامل الثاني عشر مفردات أرقامها (5 ، 8 ، 13 ، 15 ، 18، 19، 25 ، 30 ، 32 ، 33) وكلها تشبعات موجبة .

3. تشبع بالعامل الثالث خمس مفردات أرقامها (2 ، 5 ، 6 ، 19 ، 33) وكلها تشبعات موجبة ما عدا المفردة (33) فتشبعها سالب .

4. تشبع بالعامل الرابع سبع مفردات أرقامها (1 ، 6 ، 7 ، 10 ، 11 ، 33 ، 35) معظمها تشبعات موجبة ماعدا المفردتين (7 ، 11) فتشبعها سالب .

5. تشبع بالعامل الخامس خمس مفردات أرقامها (3 ، 5 ، 15 ، 18 ، 32) ومعظمها تشبعات موجبة ماعدا المفردتين (3، 18) فتشبعها سالب .

6. تشبع بالعامل السادس ثلاث مفردات أرقامها (18 ، 21 ، 28) وكلها تشبعات موجبة .

7. تشبع بالعامل السابع خمس مفردات أرقامها (7، 13، 14 ، 15 ، 35)وكلها تشبعات موجبة ما عدا المفردة (15) تشبعها سالب .

8. تشبع بالعامل الثامن سبع مفردات أرقامها (2 ، 4، 16 ، 27 ، 30 ، 31، 35) ومعظمها تشبعات سالبة ما عدا المفردات (16 ، 27، 31) فتشبعاتها موجبة.

9. تشبع بالعامل التاسع مفردتين هما (7 ، 8) تشبعا سالبا .

10. تشبع بالعامل العاشر اربع مفردات ارقامها (5، 15، 17، 28) منها المفردات (5، 17)تشبعهما موجب بينما (15 ، 28) تشبعهما سالب .

11. تشبع بالعامل الحادي عشر ثلاث مفردات ارقامها (5، 11، 26)معظمها تشبعات موجبة ماعدا المفردة (5) فتشبعها سالب .

12. تشبع بالعامل الثاني عشر اربع مفردتين ارقامها (4، 34) وكلها تشبعات موجبة.

13. تشبع بالعامل الثالث عشر اربع مفردات أرقامها (1، 7، 15، 36) وكلها تشبعات موجبة.

14. لم يتشبع بالعامل الرابع عشر اي مفردة.

15. تشبع بالعامل الخامس عشر مفردتين هما (13، 16) تشبعا موجبا ضعيفا .

وللحصول علي حل نهائي قابل للتفسير ، تم تدوير العوامل الخمسة عشر تدويرا متعامدا بطريقة الفاريماكس Varimax Orthogonal Rotation ، وهذا التدوير مكن الباحثين من افتراض أنه يمكن إيجاد حل عاملي ذي معني من خلال ستة أو سبعة عوامل ،فأجري التحليل العاملي مرة أخري بافتراض أن عدد العوامل هو ستة ، ثم بافتراض أن عدد العوامل سبعة ، وفي كل مرة تم تدوير المحاور بطريقة الفاريماكس .

وكان الحل الأخير لغرض التشخيص من حيث إمكانية الحصول علي عامل سابع بقيمة كامنة أكبر أو تساوي واحدا ، ولم يتم الحصول علي هذا العامل ، واكتفي الباحث بحل عاملي ذي ستة عوامل ، هذه العوامل الستة أعطت تفسيرا مقبولا يحقق أبسط بناء عاملي ، ويوضح ذلك الجدول التالي .

جدول رقم (11)

تشبعات مقياس الحكمة الاختيارية علي العوامل الستة بعد التدوير باستخدام المكونات الأساسية المتعاقبة .

م	العوامل					
	1	2	3	4	5	6
1	0.01 -	0.07	0.04	0.01 -	0.16 -	0.11 -
2	0.012 -	0.22 -	0.22	0.002	0.10	0.05
3	0.02	0.001	0.09	0.04 -	0.04	0.01 -
4	0.04 -	0.05 -	0.55	0.02 -	0.07 -	0.02 -
5	0.03	0.04 -	0.05 -	0.77	0.12	0.09
6	0.16	0.30 -	0.4	0.12	0.34 -	0.03 -
7	0.08 -	0.13 -	0.09 -	0.01 -	0.04	007.
8	0.10	0.02	0.30	0.13	0.004 -	0.081
9	0.02	0.16 -	0.44 -	0.05	0.011 -	0.08 -
10	0.10	0.01	0.04 -	0.01	0.56	0.02
11	0.15 -	0.16 -	0.11	0.05	0.06	0.06
12	0.24	0.02 -	0.06	0.010	0.04 -	0.06
13	0.06	0.03	0.02	0.08	0.09 -	0.02
14	0.04 -	0.65	0.10	0.01	0.11 -	0.01 -
15	0.14	0.07 -	0.05 -	0.11	0.05 -	0.11
16	0.02 -	0.01 -	0.01 -	0.04	0.01	0.09

0.08	0.10	0.06	0.02	0.12	0.02	17
006.	0.55	0.07 -	0.10	0.18	0.18	18
0.16	0.05	0.21	0.46	0.13	0.09	19
0.08	0.08 -	0.02 -	0.04	0.01	0.07	20
0.05 -	007.	0.06	0.04	0.02	0.04	21
0.05	0.02	0.11 -	0.25	0.10	0.12 -	22
0.01 -	036.	0.12	0.09	0.01 -	0.02 -	23
0.01 -	0.08	0.03	0.24	0.12	0.05	24
0.02	0.08 -	0.09	0.06	0.16 -	0.52	25
0.04 -	0.00	0.04	0.01 -	0.02 -	0.01	26
0.2	0.08	0.03	0.15	0.03	- 0.010	27
0.19	0.16	0.16 -	0.06 -	0.32	0.01	28
0.01 -	0.23	0.01 -	0.10 -	0.35	0.08 -	29
0.08	0.19	0.20	0.33	0.06	0.05	30
0.43	0.01	0.03 -	0.07 -	0.01 -	0.05	31
0.03 -	0.18	0.057	0.12	0.05	0.08 -	32
0.13	0.01 -	0.11 -	0.10 -	0.04 -	0.88	33
0.01 -	0.01	0.18	0.10	0.07 -	0.05	34
0.09 -	0.04 -	0.01	0.16	0.02	0.33	35
0.24 -	0.09	0.07 -	0.04	0.02 -	0.01 -	36

ويتضح من مصفوفة العوامل بعد التدوير ما يلي :-

أن العوامل الستة طائفية .

العامل الأول :-

وقد تشبع بهذا العامل (3) مفردات تشبعا موجبا ، وفي ضوء ما تقيسه هذه المفردات نجد أن مفردتين منهما حصلنا علي أعلي تشبع وهو (0.88 ، 0.52) يقيسان الاستفادة من معلومات محتوي مشابه للاختبار في إجابة المفردات وينتمي هذا العنصر إلي الاستدلال الاستنباطي ، ويمكن أن نطلق علي هذا العامل " استفادة المفحوص من معلومات محتوي مشابه للاختبار في الإجابة علي مفردات الاختبار المقدم له " وهو يستوعب (5.90) من التباين الكلي .

العامل الثاني :-

وقد تشبع بهذا العامل خمس مفردات تشبعا موجبا ما عدا مفردة واحدة تشبعهما سالب . وحين ننظر إلي مضمون هذه المفردة نلاحظ أنها تشير إلي استبعاد الاختيارات التي من الواضح أنها غير صحيحة والاختيار من بين الاختيارات المتبقية ويطلق علي هذا العنصر استبعاد الاختيارات غير المقبولة ، وذلك يتبع الاستدلال الاستنباطي ، كما تشير إلي استخدام التخمين وتجنب الخطأ ، ويمكن أن طلق علي هذه العامل " استبعاد الاختيارات التي من الواضح أنها غير صحيحة والاختيار أو التخمين من بين الاختيارات المتبقية " وهو يستوعب (5.4) من التباين الكلي .

العامل الثالث :-

وقد تشبع بهذا العامل خمس مفردات معظمها تشبعات موجبة ما عدا مفردة واحدة تشبعها سالب ، وحين تنظر إلي مضمون هذه المفردات نلاحظ أنها تشير إلي تجنب الخطأ في صورة قراءة التعليمات بعناية واستخدام التخمين والتحديات والتحديدات الخاصة والاختيارات المتشابهة في مقابل جعل الإجابة الصحيحة متناسقة لغويا مع الجذع ، وبذلك فقد تضمن هذا العامل عناصر خاصة

بالمفحوص ومستقلة عن مصمم وهدف الاختبار في مقابل عناصر معتمدة علي المصمم والهدف للاختبار، ويمكن أن نطلق علي هذا العامل " تجنب الخطأ في مقابل جعل الإجابة الصحيحة متناسقة لغويا مع الجذع " وهو يستوعب(4.10) من التباين الكلي ، ويعتبر عامل قطبي .

العامل الرابع :

وقد تشبع بهذا العامل مفردتان تشبعا موجبا ، وحين ننظر إلي مضمون كل منهما نلاحظ أن أحداهما تشير إلي استبعاد الاختيارات المتشابهة والأخرى تشير إلي الاختيارات المتداخلة ، ويمكن أن نطلق علي هذا العامل " استخدام الاختيارات المتشابهة والمتداخلة " وهو يستوعب (3.60) من التباين الكلي .

العامل الخامس :

وقد تشبع بهذا العامل أربع مفردات ، منها تشبعات موجبة (مفردتان) / وحين ننظر إلي مضمون هذه المفردات نلاحظ أنها تشير إلي التعرف علي التحديدات الخاصة المرتبطة بالاختيارات الصحيحة وغير الصحيحة وهذا العنصر ينتمي إلي إستراتيجية استخدام دليل أو رمز ، ومراعاة القصد أو الهدف للمفردة والأخذ في الاعتبار مادة الموضوع وصعوبة المفردات المجاورة عند تفسير وإجابة المفردة المعطاة للمفحوص ، ويمكن أن نطلق علي هذا العامل وهو عامل قطبي " التعرف علي التحديدات الخاصة المرتبطة بالاختيارات الصحيحة وغير الصحيحة في مقابل مراعاة القصد أو الهدف للمفردة المقدمة للمفحوص " ، وهو يستوعب (3.30) من التباين الكلي .

العامل السادس :

وقد تشبع بهذا العامل مفردتان تشبعا موجبا ، وحين ننظر إلي مضمون كل منهما نلاحظ أن أحدهما تشير إلي الاختيارات المتشابهة والأخرى تشير إلي الاستفادة من الزمن المتبقي في مراجعة الإجابات وهذا العنصر ينتمي إلي إستراتيجية استخدام الوقت ، ويمكن أن نطلق علي هذا العامل " استخدام الوقت " ، وهو يستوعب (3.00) من التباين الكلي .

لقد قدم التحليل العملي لمقياس الحكمة الاختبارية دليلا قويا علي مدي كفاءة مفردات هذا المقياس وصلاحيتها في قياس الجوانب التي صممت من أجلها ، فنجد سواء في التحليل الأول أو الثاني أن معظم المفردات علي العوامل ذات تشبعات دالة مرتفعة .

سادسا : إعداد مقياس الحكمة الاختبارية في صورته النهائية :

لقد شمل التعديل الذي أدخل علي المقياس نتيجة لخطوات التقنين كراسة الأسئلة وورقة الإجابة بالنسبة لكراسة الأسئلة بناء علي صدق وثبات المفردات السابق بيانه ، فقد تم حذف أحد عشر سؤالا من مقياس الحكمة الاختبارية ارقامها (2 ، 3 ، 6 ، 6 ، 9 ، 17 ، 21 ، 23 ، 27 ، 28 ، 29 ، 36) ، وبذلك اصبح عدد الأسئلة المكونة للمقياس في الصورة النهائية (25) سؤالا (بدلا من 36 سؤالا) .

ونظرا لحذف عدد من الأسئلة (أحد عشر) فقد بين الباحث عدد الأسئلة لكل إستراتيجية مرة ثانية بعد هذا التعديل كما يلي :

- إستراتيجية استخدام دليل أو رمز : وتضم (7) أسئلة ارقامها : 8 ، 11 ، 13 ، 19 ، 20 ، 22 ، 24
- إستراتيجية الاستدلال الاستنباطي : وتشمل (7) اسئلة ارقامها : 3، 5 ، 7 ، 10 ، 15 ، 18 ، 23

- إستراتيجية استخدام الوقت : وتضم (4) أسئلة أرقامها : 1 ، 9 ، 17 ، 21
- إستراتيجية تجنب الخطأ : سؤالان هما : 2 ، 12
- إستراتيجية مراعاة القصد أو الهدف : وتضم (3) أسئلة أرقامها : 6 ، 16 ، 25

أما ورقة الإجابة فقد أعيد تنظيمها بحيث أصبح عدد الإجابات يساوي (25) إجابة ، بدلا من (36) إجابة .

سابعا : الزمن المناسب للمقياس :

يعد المقياس من الاختبارات الموقوتة – أي التي لها زمن محدد : وتتأثر درجات الاختبارات الموقوتة تأثرا مباشرا بزمن الإجابة ، وبذلك تصبح مشكلة تحديد الزمن من أهم المشاكل العلمية التي يواجهها الباحث في إعداده للاختبارات الجديدة . وقد كشف " فؤاد البهي السيد " عن المعادلة الرياضية التي تحدد العلاقة القائمة بين متوسطات الدرجات والأزمنة ، ومعاملات السهولة ، وأحدي الصور الرياضية لتلك المعادلة هي :

$$ز_2 = \frac{م_2}{م_1} \times ز_1$$

حيث يدل :

الرمز ز2 علي الزمن المناسب للاختبار ، ز1 علي الزمن التجريبي للاختبار ، والرمز م2 علي المتوسط المرتقب للدرجات ، والرمز م1 علي المتوسط التجريبي للدرجات .

وباستخدام هذه الطريقة وجد أن الزمن المناسب للمقياس هو (14.5) دقيقة .

ثامنا : المعايير :

تدل المعايير علي الأداء الاختباري للعينة ، وبذلك يمكن تحويل الدرجة الخام إلي نوع من المقياس النسبي . كما أن القيمة الخام من اي مجموعة من القيم لا تعطي معني أو دلالة .

ولقد تضمن حساب المعايير امرين هامين هما : تقدير درجة المفحوص ، وتفسير الدرجة الخام

أ- تقدير درجة المفحوص :-

اتبع الباحث في تقدير الدرجة مفتاح التصحيح من النوع المثقب ، وهو من الورق المقوي بحجم ورقة الإجابة ومثقبة بعدد أسئلة المقياس ، ويوضح أمام رقم كل سؤال درجات الإجابات الصحيحة ، وحين يوضع المفتاح علي ورقة الإجابة يظهر امام كل سؤال رمز الإجابة التي اختارها المفحوص ، وبذلك يأخذ كل سؤال درجة الرمز الموضوع أمامه ، ورموز الاختيارات هي (ا ، ب ، ج ، د) ، كما يوجد ثقب كبير في نهاية المفتاح ترصد به الدرجة الكلية للمفحوص الناتجة عن جمع درجات الأسئلة المكونة للمقياس .

ب- تفسير الدرجة الخام :-

لمعرفة دلالة اي درجة من الدرجات الخام التي يعطيها المقياس قام الباحث بتفسير الدرجات الخام التي حصل عليها من تطبيق المقياس علي عينة البحث الأصلية (2148) طالبا وطالبة . واستخدام طريقة التفسير بواسطة المئينيات حيث هي النقط التي تقسم التوزيع التكراري إلي أجزاء مئوية . ويوضح الجدول رقم (12) الرتب المئينية المقابلة للدرجات الخام

جدول رقم (12)

الرتب المئينية المقابلة للدرجات الخام

الدرجة الخام	الرتبة المئينية	الدرجة الخام	الرتبة المئينية
15	1	27	38
16	2	28	47
17	3	29	57
18	4	30	64
19	6	31	77
20	8	32	84
21	10	33	90
22	14	34	94
23	18	35	97
24	24	36	99
25	31		

مقياس الحكمة الاختبارية

" الصورة الأولية "

إعداد

دكتور / فتحي عبد الحميد عبد القادر

تعليمات :

1. يستخدم هذا المقياس في قياس الحكمة الاختبارية لدي الطلبة والطالبات .
2. يتكون المقياس من (36) سؤالاً يلي كل سؤال أربعة اختيارات رموزها (أ ، ب ، جـ ، د) .
3. علي الطالب / الطالبة اختيار الإجابة الصحيحة من وجهة نظره / نظرها ثم وضع رمز الإجابة في المستطيل الموجود علي يسار رقم السؤال في ورقة الإجابة .
4. لا تكتب شيئاً في كراسة الأسئلة واكتب في ورقة الإجابة فقط .
5. دون الوقت الذي تستغرقه في الإجابة علي المقياس في ورقة الإجابة بعد الانتهاء من الإجابة علي المقياس .

مع الشكر ،،،،

1- عند تناولك الاختبار فانك تفضل :

أ- الإجابة ببطء .

ب- الإجابة بسرعة ؟.

ج- الإجابة بسرعة مع مراعاة الإتقان .

د- الإجابة بسرعة بغض النظر عن النتيجة .

2- عندما تضاف بلورات من مادة كيميائية للماء فانه :

أ- يعطي حرارة .

ب- ترتفع درجة حرارته .

ج- يتحول المحلول للون الأزرق .

د- يصبح المحلول دافئاً .

3- الهيئة التي لها سلطة علي الطيران المدني سميت :

أ- هيئة علم الطيران المدني .

ب- لجنة فحص الخطوط الجوية .

ج- توزيع المسافرين علي الطرق الجوية .

د- لجنة الطيران .

4- عند تناول الاختبار فانك تفضل :

أ- قراءة التعليمات بعناية وتحديد المطلوب في الإجابة .

ب- عدم التركيز علي التعليمات .

ج- قراءة التعليمات قراءة سطحية .

د- البدء في الإجابة دون النظر لتعليمات .

5- الاضطراب العقلي الذي صنف علي اضطراب عصبي هو :

أ- هستريا .

ب- الفصام .

ج- كآبة سن اليأس .

د- انشطار الشخصية .

6- التعديلات التي حدثت للدستور تسمي :

أ- تعديلات تنظيم جديد .

ب- تعديلات رئاسية .

ج- رسالة في العدل ـتعديلات تشريعية .

7- عند إجابتك لأسئلة الاختبار ذي الاختيار من متعدد فانك :

أ- تخمن إذا لم يوجد وقت .

ب- تخمن في كل الحالات الاختبارية .

ج- لا تخمن مهما كان الدافع .

د- تخمن حتي لو وجد وقت .

8- عند إتحاد مادة كيميائية مع الأكسجين فأنه :

أ- يعطي لوناً احمر .

ب- تنطلق حرارته .

ج- يعطي لوناً اصفر .

د- يعطي لوناً برتقالياً .

9- كتاب طه حسين الأيام مثال لـ:

أ- ترجمة حياة شخص .

ب- رواية علمية .

ج- كتاب سياسي .

د- قصة حياة الكاتب بقلمه .

10- الضوء يسير علي هيئة :

أ- خط مستقيم .

ب- موجات .

ج- قطع ناقص .

د- دوائر .

11- اي من الأتي تعتبر ميزة لاستخدام الأسرة العالية في المستشفيات :

أ- تكلفة الأسرة العالية اكثر من الأسرة العادية .

ب- رعاية المرضي تكون سهلة .

ج- احتمال سقوط امرضي من الأسرة العالية .

د- الأسرة العالية لا تستخدم طوال اليوم .

12- جيمس وايت عرف جيداً لأنه :

أ- اول مدرس اسود .

ب- كان محافظاً (رئيس بلدية) مرتين .

ج- رجل إطفاء جيد .

د- كان ملاكماً ماهراً .

13-	اثناء اداء الاختبار فانك . أ- تحدد زمناً خاص لكل سؤال . ب- تحدد زمناً متساوياً لجميع الأسئلة بغض النظر عن سهولتها او صعوبتها . ج- تجيب عن الأسئلة دون النظر للزمن . د- تسير بمعدل زمني يتناسب مع المطلوب بعد كل فترة زمنية معينة .
14-	الدولة التي بها سكان اكثر هي : أ- السودان . ب- نيويورك . ج- ج . م . ع . د- باريس .
15-	اشتهر قدماء المصريين لأنهم : أ- أول من مارس فن التحنيط . ب- صنعوا شباك صيد السمك . ج- اصطادوا الحيوانات من الغابات . د- اهتموا بالصناعة .
16-	عند اداء الاختبار فانك تفضل : أ- قراءة المفردات بعناية وتحديد مطلب السؤال قبل الإجابة عليه . ب- قراءة المفردات ومطلبها . ج- عدم التركيز علي القراءة
17-	المحافظة التي تضم عدد سكن اكثر في ج . م . ع هي : أ- الشرقية . ب- طنطا . ج- القاهرة . د- المنصورة .
18-	ابن سينا مشهور في : أ- علم الطب .

ب- تطوير جميع المقاييس الموسيقية في العام .

ج- تأليف جميع الألحان الموسيقية .

د- ترجمة جميع الدراسات الإنجليزية .

19- عند استخدامك للتخمين في اختيار الإجابة الصحيحة فأنك :

أ- تختار الإجابة عشوائياً .

ب- تأخذ في الاعتبار محتوي الاختيارات .

ج- تهمل محتوي الاختيارات .

د- تستبعد الإجابة الخطأ ثم تخمن من الإجابات المتبقية .

20- في تعديلات الدستور :

أ- منع صناعة ورواج المشروبات المسكرة .

ب- الغي المنع أو التحريم لصناعة ورواج المشروبات المسكرة .

ج- أعطيت المرأة الحق في الانتخاب .

د- منع رئيس الجمهورية من الانتخاب للمنصب أكثر من مرتين .

21- عرفت قارة افريقيا بسبب :

أ- وجود حيوانات في الغابات ،

ب- عدم وجود أفراد من كبار السن يعيشون فيها .

ج- صناعة السفن .

د- جميع الناس فيها مزارعون .

22- يتحرك الإلكترون في :

أ- الذرة .

ب- المدارات .

ج- المدار المتفق مع طاقة الإلكترون .

23- اشتهر جاليلو بسبب :

أ- تأليف العديد من الكتب .

ب- تقديم نقود لإنشاء المدارس .

ج- تطعيم الحيوانات .

د- اختراع التلسكوب الفلكي .

24- عند اجابك لأسلة الاختبار فانك :
أ- تهمل أو تؤجل المفردات التي تحتاج زمناً أكثر .
ب- تخمن إجابة المفردات التي تستغرق زمناً أكثر .
ج- تبدأ بإجابة المفردات التي تحتاج إلي زمن أكثر .
د- تجيب علي المفردات المتأكد من صحتها ثم تخمن إجابة المفردات الاخري .

25- اي من الأربعة الآتية ثابت الحرارة :
أ- الضفدعة .
ب- العصفور .
ج- الثعبان .
د- السحلية .

26- الحزب الوطني الديمقراطي مركزه الرئيسي في :
أ- مدينة الإسكندرية .
ب- مدينة القاهرة .
ج- مدينة أسوان .
د- مدينة الزقازيق .

27- بعد إجابة اختبار فانك :
أ- تراجع جميع الإجابات .
ب- تراجع الإجابات غير المتأكد منها .
ج- تراجع إجابات المفردات الصعبة .
د- تهمل مراجعة الإجابات .

28- يساعد إيراد دخل قناة السويس في :
أ- بناء البنوك .
ب- بناء المستشفيات .
ج- بناء المدارس .
د- زيادة الدخل القومي لمصر .

29- عندما تلجأ للتخمين في اختيار الإجابة الصحيحة فانك :

أ- تحدد جميع الاحتمالات الإجابة الصحيحة .

ب- لا تستبعد أي إجابة قبل التخمين .

ج- لا تهتم بتحديد الاحتمالات الإجابة الصحيحة .

د- لا تختر الإجابة علبي اساس محتوي البدائل .

30- تتميز اشعة جاما :

أ- بأنها موجات كهرومغناطيسية .

ب- بان كتلتها كبيرة .

ج- بأن كتلتها متوسطة .

د- بأن كتلتها صغيرة .

31- إذا تناولت اختبارا فأنك :

أ- توفر جزءاً من زمن الاختبار للمراجعة .

ب- لا توفر جزءاً من الزمن المخصص للاختبار .

ج- تستفيد من الزمن المتبقي بعد تكملة الاختبار في مراعاة الإجابات .

32- عدد سكان ج .م .ع أكثر من :

أ- أربعين مليوناً .

ب- خمسين مليوناً .

ج- اثنين وخمسين مليوناً .

د- خمسة وخمسين مليوناً .

33- أي من الأربعة حيوانات الآتية متغير الحرارة :

أ- الثعبان .

ب- الكلب .

ج- الفيل .

د- الحوت .

34. عدد الكتب في مدرستنا يقل عن :

أ- مائة الف كتاب .

ب- تسعين الف كتاب .

ج- ثمانين الف كتاب .

د- سبعين ألف كتاب .

35. يوجد البروتون في :

أ- الجزئي .

ب- الذرة .

ج- النواة .

د- المدارات .

36. مر علي انشاء الجمعية العامة بالأمم المتحدة فترة تقل عن:

أ- خمس وأربعين سنة .

ب- أربعين سنة .

ج- خمس وثلاثين سنة .

د- ثلاثين سنة .

مقياس الحكمة الاختبارية

" الصورة النهائية "

تعليمات :

1. يستخدم هذا المقياس في قياس الحكمة الاختبارية لدي الطلبة والطالبات .
2. يتكون المقياس من (25) سؤالاً يلي كل سؤال أربعة اختيارات رموزها (أ ، ب ، جـ ، د) .
3. علي الطالب / الطالبة اختيار الإجابة الصحيحة من وجهة نظره / نظرها ثم وضع رمز هذه الإجابة في المستطيل الموجود علي يسار رقم السؤال في ورقة الإجابة .
4. لا تكتب شيئاً في كراسة الأسئلة واكتب في ورقة الإجابة فقط .
5. الزمن المخصص لهذا الاختبار (15) دقيقة .

مع الشكر ،،،،،،،

1- عند تناولك الاختبار فإنك تفضل: أ- الإجابة ببطء . ب- الإجابة سرعة . ج- الإجابة بسرعة مع مراعاة الإتقان . د- الإجابة بسرعة بغض النظر عن النتيجة .
2- عند تناول الاختبار فإنك تفضل : أ- قراءة بعناية وتحيد المطلوب في الإجابة . ب- عدم التركيز علي التعليمات . ج- قراءة التعليمات قراءة سطحية .

د- البدء في الإجابة دون النظر للتعليمات .

3- الاضطراب العقلي الذي صنف علي الاضطراب عصبي هو :

أ- هستريا .

ب- الفصام .

ج- كآبة سن اليأس .

د- انشطار الشخصية .

4- عند إجابتك لأسئلة الاختبار ذي الاختيار من متعدد فانك :

أ- تخمن إذا لم يوجد وقت .

ب- تخمن في كل الحالات الاختبارية .

ج- لا تخمن مهما كان الدافع .

د- تخمن حتي لو وجد وقت .

5- عند إتحاد عنصر من الأكسجين فأنه :

أ- يعطي لوناً أحمر.

ب- تنطلق حرارته .

ج- يعطي لوناً أصفر .

د- يعطي لوناً برتقالياً .

6- الضوء يسير علي هيئة :

أ- خط مستقيم .

ب- موجات .

ج- قطع ناقص .

د- دوائر .

7- اي من الأتي تعتبر ميزة لاستخدام الأسرة العالية في المستشفيات :

أ- تكلفة الأسرة العالية اكثر من الأسرة العادية .

ب- رعاية المرضي تكون سهلة عندما تكون الأسرة عالية .

ج- احتمال سقوط المرضي من الأسرة العالية .

د- الأسرة العالية لا تستخدم طوال اليوم .

8- جيمس وايت عرف جيداً لأنه :

أ- اول مدرس أسود .

ب- كان محافظاً (رئيس بلدية) مرتين .

ج- رجل إطفاء جيد .

د- كان ملاكماً ماهراً .

9- اثناء اداء الاختبار فإنك :

أ- تحدد زمناً خاص لكل سؤال .

ب- تحدد زمناً متساوياً لجميع الأسئلة بغض النظر عن سهولتها أو صعوبتها .

ج- تجيب علي الأسئلة دون النظر للزمن .

د- تسير بمعدل زمني يتناسب مع المطلوب .

10- الدولة التي بها سكان اكثر هي :

أ- السودان .

ب- نيويورك .

ج- ج . م . ع .

د- باريس .

11- اشتهر قدماء المصريين لأنهم :

أ- أول من مارس ن التحنيط .

ب- صنعوا شباك صيد السمك .

ج- اصطادوا الحيوانات من الغابات

د- اهتموا بالصناعة .

12- عند اداء الاختبار فإنك تفضل .

أ- قراءة المفردات بعناية وتحديد مطلب السؤال قبل الإجابة عليه .

ب- قراءة المفردات ومطلبها .

ج- عدم التركيز علي القراءة .

13- ابن سينا مشهور في :

أ- علم الطب .

ب- تطوير جميع المقاييس الموسيقية في العالم

ج- تأليف جميع الألحان الموسيقية . د- ترجمة جميع الدراسات الإنجليزية .
14- عند استخدامك للتخمين في اختيار . الإجابة الصحيحة فأنك : أ- تختار الإجابة عشوائياً . ب- تأخذ في عين الاعتبار محتوي الاختيارات . ج- تهمل محتوي الاختيارات . د- تستبعد الإجابة الخطأ ثم تخمن من الإجابات المتبقية .
15- في تعديلات الدستور : أ- منع صناعة ورواج المشروبات المسكرة. ب- الغي المنع أو التحريم لصناعة ورواج المشروبات المسكرة . ج- أعطيت المراة الحق في الانتخاب . د- منع رئيس الجمهورية من الانتخاب للمنصب مرتين .
16- يتحرك الإلكترون في : أ- الذرة . ب- المدارات . ج- المدار المتفق مع طاقة الإلكترون .
17- عند إجابتك لأسئلة الاختبار فانك : أ- تهمل أو تؤجل المفردات التي تحتاج زمناً أكثر . ب- تخمن إجابة المفردات التي تستغرق زمناً أكثر . ج- تبدأ بإجابة المفردات التي تستغرق إلي زمن أكثر . د- تجيب علي المفردات المتأكد من صحتها ثم تخمن إجابة المفردات الاخري .
18- اي من الأربعة الآتية ثابت الحرارة : أ- الضفدعة . ب- العصفور . ج- الثعبان . د- السحلية .

19- الحزب الوطني الديمقراطي مركزه الرئيسي في :

أ- مدينة الإسكندرية .

ب- مدينة القاهرة .

ج- مدينة أسوان .

د- مدينة الزقازيق .

20- تتميز اشعة جاما :

أ- بأنها موجات كهرومغناطيسية .

ب- بأن كتلتها كبيرة .

ج- بأن كتلتها متوسطة .

د- بأن كتلتها صغيرة .

21- إذا تناولت اختباراً فإنك :

أ- توفر جزءاً من زمن الاختبار للمراجعة .

ب- لا توفر جزءاً من الزمن المخصص للاختبار .

ج- تستفيد من الزمن المتبقي بعد تكملة الاختبار في مراجعة الإجابات .

22- عدد سكان ج . م . ع اكثر من :

أ- أربعين مليون .

ب- خمسين مليون .

ج- اثنين وخمسين مليون .

د- خمسة وخمسين مليون .

23- أي من أربعة الحيوانات الآتية متغيرة الحرارة :

أ- الثعبان .

ب- الكلب .

ج- الفيل .

د- الحوت .

24- عدد الكتب في مدرستنا يقل عن :

أ- مائة الف كتاب .

ب- تسعين الف كتاب .

ج- ثمانين الف كتاب .

د- سبعين الف كتاب .

25- يوجد البروتون في :

أ- الجزئي .

ب- الذرة .

ج- النواة .

د- المدارات .

" ورقة إجابة مقياس الحكمة الاختبارية "

اسم الطالب / ....   السنة الدراسية : ....

النوع : ....   العمر : ....

رقم السؤال	رمز الإجابة	رقم السؤال	رمز الإجابة	رقم السؤال	رمز الإجابة
1		10		19	
2		11		20	
3		12		21	
4		13		22	
5		14		23	
6		15		24	
7		16		25	
8		17		الدرجة الكلية	
9		18			

الباب الثاني

الاختبارات النفسية

الفصل الخامس

نشأة الاختبارات النفسية (منظور تاريخى)

أى نظرة عامة قصيرة على الاسبقيات التاريخية وأصول الاختبار النفسى يجد أنها تساعد فى فهم اختبارات الوقت الحاضر. فالاتجاه فى أى اختبار نفسى معاصر متقدم يمكن أن يوضح فى ضوء بوادر مثل هذه الاختبارات والتعقيدات الخاصة بالإضافة إلى الفوائد التى تميز الاختبارات الحالية لتصبح أكثر وضوحا على نفس النمط عندما ننظر إلى تلك الخلفية التى وجدوا فيها هذه الاختبارات ، فهذا الفصل يهتم فقط بالأسبقيات والتطور المبكر لحركة الاختبار ككل. فهناك تطورات حديثة أكثر مناقشة فى الفصول التالية بالارتباط بأنواع معينة من الاختبارات مثل اختبارات القدرة ، وأهمية الاختبارت ، بالإضافة إلى استعمال الاختبار فى مجالات تعليمية وصناعية وطبية واستشارية ، لقد فقدت أصول الاختبار فى العصور القديمة حيث كانت هناك حسابات متكررة تسود نظام فحص الخدمة المدنية فى الإمبراطورية الصينية لحوالى 2000 سنة (بومان 1989) كان بين اللغة اليونانية القديمة اختبار ملحقا بالعملية التربوية ، حيث كانت الاختبارات تستعمل لتقييم الإجادة الطبيعية بالإضافة إلى المهارات الثقافية (دويل 1970) فمنذ العصور الوسطى نجد الجامعات الأوروبية قد اعتمدت على فحوص رسمية فى منح الشهادات ودرجات الشرف على أية حال ، لن نحتاج أن نذهب إلى أبعد من القرن التاسع عشر لتميز التطورات الرئيسية فى الاختبارات المعاصرة.

أهمية التصنيف المبكر والتدريب للأفراد المتخلفون عقليا:

لقد شهد القرن التاسع عشر يقظة قوية فى الاهتمام بالمعالجة الإنسانية للمعاقين عقليا والأشخاص المجانين ، فكان قبل ذلك هناك إهمال ، سخرية ،

وتعذيب لمثل هذه الشريحة من الأفراد نتيجة القلق المتزايد بالعناية الصحيحة بالمشاكل العقلية للناس جاءت إدراك بعض المعايير الموحدة لتمييز وتصنف هذه الحالات فتطلب العديد من المؤسسات الاجتماعية للعناية بالأشخاص المتخلفون عقليا فى كلتا من أوروبا والولايات المتحدة حيث عملتا على البدء فى إدخال معايير ونظام موضوعي.

نلاحظ أن التصنيف السريع ضروري بالدرجة الأولى فى التفرقة بين الشخص المصاب الجنون والشخص المتخلف عقلي. وقد أوضح (أسكوول) أن الاضطرابات الانفعالية يمكن أن تكون مصحوبة بتدهور ثقافى فى المستوى الطبيعى الأولى. ونجد أن هذا المستوى قد تم تميزه فى الأساس عن طريق العيب العقلى من قبل الأب أو فى مرحلة الطفولة المبكرة ، وقد تم ذكر ذلك فى مجلد نشر عام 1838 عن طريق الطبيب الفرنسى (أسكوول) والتى تتفاوت ما بين الحالة الطبيعية حتى البلاهة والتى تمثل الدرجة المنخفضة ، كذلك فقد قام (أسكوول) بعدة إجراءات استنتج منها أن استعمال الفرد للغة تزودنا بالمعيار الذى يمكن أن نعتمد عليه فى ارتقاء اللغة والمستوى الثقافى للفرد.

العلماء النفسانيون التجريبيون (الأوائل):

لم يعتنى علماء النفس الأوائل فى القرن الـ 19 بشكل عام بقياس الفروق الفردية ، وكان هدفهم فى هذه الفترة هو صياغة التوصيفات العامة للسلوك البشرى ، وكان التركيز على المشتركات أكثر من الاختلافات فى السلوك. الفروق الفردية إما أن يتم تجاهلها أو قبولها كضرورة تحدد مدى ملائمة التعميمات هكذا ، أن حقيقة أن الفرد يستجيب بشكل مختلف عند ملاحظة فى ظروف مماثلة يؤخذ على الاعتبار أنه شكل من اشكال الخطأ ، وجود مثل هذا الخطأ أو تغير الفرد جعل التعميمات متقاربة هذا كان الاتجاه نحو الفروق

الفردية والتى سادت فى مثل المختبرات مثل الذى كان الاتجاه نحو الفروق الفردية والتى أسست عن طريق ونت 1879 حيث تلقى العلماء النفسيون التجريبيون الأوائل تدريبهم.

فى اختيار موضوعاتهم ، وفى مراحل كثيرة من علمهم ، عكس مؤسسى علم النفس التجريبى تأثير خلفياتهم فى علم الفيزياء والفسيولوجيا ، المشاكل التى درست فى مختبراتهم أعتنت بشكل كبير بالحساسية للحوافز البصرية والمسعية وغيرها هذا التأكيد على الظواهر الأحساسية كانت تباعا يعكس طبيعة اختبارات النفسية الأولى.

ورغم ذلك ربما نلاحظ طريق آخر والذى فيه أثر العلماء الأوائل على تأسيس حركة الاختبار. العلماء الأوائل وضحوا الحاجة إلى التحكم الشديد فى الظروف التى تتم تحتها الملاحظات ، على سبيل المثال ، صياغة التوجهات التى تعطى للمشارك ربما تزيد أو تقلل سرعة استجابة الفرد ، ألوان المجال المحيط ربما يعدل ظهور الحوافز الحسية. أهمية صنع الملاحظات على المشاركين تحت ظروف موحدة يكون واضح. مثل هذا التوحيد فى الإجراءات أصبح واحدة من العلامات الخاصة للاختبارات النفسية.

مساهمات فرنسيس جالتون:

العالم البيولوجى فرنسيس جالتون كان هو أول المسئول عن انطلاق حركة الاختبار ، العامل الموحد فى نشاطات البحث المتنوعة هو اهتمام بالوراثة. فى طريقة البحث على الوراثة ن أدرك جالتون الحاجة إلى قياس الصفات المرتبطة والغير مرتبطة بالأشخاص. فقط بهذا الطريقة اكتشف على سبيل المثال ، الدرجة الدقيقة للتشابهة بين الوالدين والذرية ، الأخوة ، الأخوات ، أبناء العم والخالة ، التوائم بهذه النهاية ، كان جالتون مفيدا فى البحث فى عدد من

المؤسسات التعليمية لحفظ تسجيلات منظمة فى علم الإنسان على التلاميذ ، أسس أيضا مخترعات علم الإنسان "العرض العالمى 1884" ، فى تلك المختبرات يمكن للزائرين بعد دفع نقود أن يقيسوا السمات الفيزيائية وأخذ اختبارات شدة الرؤية والسمع والقوة العضلية وغير ذلك من الوظائف الإحساس ، وعندما أغلق العرض ، المخبرات نقلت إلى متحف كينستون فى لندن ، حيث عملت لمدة 6 سنوات ، ويمثل هذه الطرق تراكمت أول أكبر قاعدة معلومات عند الفروق الفردية فى نفسية بسيطة.

جالتون نفسية ابتكر معظم الاختبارات البسيطة التى أدارها فى مختبراته ، الكثير منها مازالت مشهورة فى أشكالها الأصلية والمعدلة ، استخرج جالتون موازين لقياس التمييز الحركى انطلاقا من اعتقاده بأن اختبارات التمييز الحسى ربما يخدم كوسائل لقياس ذكاء الشخص فى هذا المجال كان جالتون متأثرا بشكل جزئى بنظريات "لوك" ، وهكذا كتب جالتون أن المعلومات الوحيدة التى تصل إلينا بخصوص الأحداث الخارجية لتمر عبر طريق الحواس ، وكلما كانت الحواس المستقبلية مختلفة كلما كان المجال الذى نحكم فيه اكبر (جالتون 1883).

لاحظ جالتون أيضا أن الأشخاص الذين لديهم إعاقة ذهنية شديدة يأتى ليكون خلل فى القدرة على التمييز بين الحر والبرد والألم ، ملاحظة تطوى اعتقاده أن القدرة على التمييز الحسى ، يكون إجمالا الأعلى بين القدرات الذكائية (جالتون 1883) جالتون ايضا هو الرائد فى تطبيق ميزان التصنيف وأساليب الاستفتاء وأيضا فى استخدام أسلوب "الاتحاد الحر" الذى وظف لأغراض عديدة ، مساهمات أخرى جالتون نجدها فى تطويره للطرق الإحصائية فى تحليل المعلومات على الفروق الفردية. اختار جالتون عدد من الأساليب مستمدة من علماء الرياضيات. تلك الأساليب وضعا فى مثل هذا الشكل ليسمح

باستخدامها عن طريق الباحثين الرياضيون غير المدرسين الذين ربما يتعاملون مع نتائج الاختبارات ، وبذلك وسع استعمال الإجراءات الإحصائية لتحليل معلومات الاختبارات مرحلة عمل جالتون استمرت ونقلت عن طريق كثير من تلاميذه كان ابرزهم هو (كول بيرسون).

كاتيل والاختبارات العقلية المبكرة:

مكانة هامة و خاصة فى تطوير الاختبار النفسى شغلها العالم النفسى الأمريكى كاتيل ، علم النفس التجريبى وحركة أصبحت الاختبارات الحديثة فى عمل ماتيل ، من أجل درجة الدكتوراه ، أكمل رسالته على زمن رد الفعل تحت إشراف (وندت) ، اثناء إعطاء محاضراته فى كامبريدج عام 1888 ، دعم اهتمام كاتل بقياس الفروق الفردية باتصاله بجالتون وفى عودته إلى أمريكا كان كاتيل نشيطا فى إنشاء المختبرات للعالم النفسى التجريبى وفى توصع حركة الاختبارات.

فى عام 1890 نشر مقال لكاتيل ، فيه المصطلح "الاختبار الذهنى" استخدم لأول مرة فى مواد علم النفس ، وصف هذا المقال سلسلة من الاختبارات التى أديرت سنويا لتلاميذ الكليات لتحديد مستوى ذكائهم ، الاختبارات التى أديرت للرد تتضمن مقاييس القوة العضلية ، سرعة الحركة ، الإحساس بالألم ، حدة البصر والسمع ، تحديد الوزن ، زمن رد الفعل ، الذاكرة ومتشابة ذلك.

فى اختباره للاختبارت ، كان تفصيل كاتيل مدعم بحقيقة أن تلك الوظائف البسيطة يمكن قياسها بدقة ، بينما تطور المقاييس للوظائف الأكثر تعقيدا.

اختبارات كاتيل كانت نموذج لعدد من سلسلة الاختبارات المتطورة اثناء العقد الأخير للقرن الـ 18 . مثل هذه السلسلة من الاختبارات اديرت لأطفال

المدارس ، طلاب المدارس والبالغين المتنوعين. فى شيكانو 1893 قام (جاسترو) بإنشاء معرض والذى يقوم الزائرون المدعون ليأخذوا اختبارات الإحساس ، الحركة ، عمليات الإدراك البسيطة ومقارنة مهارتهم بالمعايير. محاولات قليلة لتقييم مثل هذه الاختبارات الميكرة انتجت نتائج مشجعة. أداء الأفراد أظهرت توافق من اختبار إلى آخر وقد عرضت أنه لا علاقة لتقدير معتمد لمستوى الذكاء معتمدا على تصنيف المدرسين أو التصنيف الأكاديمى.

عدد من سلاسل الاختبارات جمعت بواسطة علماء نفسيين أوروبيون جاءت لتغطى إلى حد ما وظائف أكثر تعقيدا. أعد (سكريبلن) الذى كان مهتما أساس بالامتحان الطبى للمرضى النفسيين ، سلسلة من الاختبارات لقياس ما اعتبره كعوامل أساسية فى توصيف الفرد ، الاختبارات ، التى وظفت أساسا العمليات الرياضية ، صممت لقياس التأثيرات الحركية ، الذاكرة ، القابلية للإرهاق ، وتشوش الذهن ، عالم نفسى آخر ألمانى قام بإدارة اختبارت الحساب الرياضى ، امتداد الذاكرة.

وفى مقال نشئ فى فرنسا عام 1895 ، نقد (هنرى ، بنت) معظم سلاسل الاختبارات المتاحة لكونها حسبة بشكل كبير والتركيز بشكل أكبر من اللازم على القدرات البسيطة الخاصة ، وناقشوا أن فى قيامى الأكثر تعقيدا ، الدقة الكبيرة ليست مطلوبة ، لأن الفروق الفردية أكبر فى تلك الوظائف ، قائمة أوسع من الاختبارات كانت مقترحة تغطى مثل تلك الوظائف مثل الذاكرة ، التخيل ، الانتباه ، الاستيعاب ، القدرة على الاقتراح وغيرها. فى تلك الاختبارات يمكننا أن نتعرف على الاتجاهات التى كانت تقود موازين الذكاء الشهيرة لـ (بنيت).

بنية وثورة اختبارات الذكاء:

بنية ومجموعة عملة خصصوا لسنوات عديدة لتنشيط البحث على طرق قياسى الذكاء ، طرق عديدة جريت وتضمنت قياس شكل الجمجمة ، الوجه واليد وتحليل الكتابة اليدوية والنتائج قادت إلى اعتقاد أن القياس المباشر للوظائف العقلية المعقدة مبشرة بالنجاح ، فى عام 1904 ، وزير التعليم عين "بنية" ليدرس الإجراء أساسى من تعليم الأطفال المعاقين ، هذا كان بالاتصال مع (جنية) بالتعاون مع (سيمون) اعدوا "ميزان بنية- سيمون".

هذا الميزان الذى عرف بميزان عام 1905 يتكون من 30 مشكلة أو اختيار رتبت فى صعوبة متصاعدة . مستوى الصعوبة يحدد بإدراة الاختبارات لخمسين طفل عادى من عمر 3 إلى 11 سنة ولأطفال بالغين معاقين ، الاختبارات صممت لتغطى تنوع واسع من الوظائف مع التأكيد على الحكم ، الاستيعاب والتفكير والذى اعتبر "بنيت" كمكونات أساسية للذكاء ، على الرغم من أن الاختبارات الحسية كانت متضمنة ن نسبة أكبر من المحتوى اللفظى كان موجودا فى الميزان بشكل اكبر مما تأثر عليه سلاسل الاختبارات فى هذا الوقت ميزان 1905 قدم كوسيلة إليه غير نهائية.

فى الميزان الثامن ، ميزان 1908 ، عدد الاختبارات زاد وتم التخلص من الاختبارات الغير مرضية من الميزان الأول ، وكل اختبارات جمعت داخل مستويات اعمار على أساس الأداء لـ 300 طفل عادى بين 3 إلى 13 سنة وهكذا فى مستوى الـ 3 سنوات تجاوزن الاختبارات 80٪ إلى 90٪ لسنوات الـ 3، وفى مستوى الأربع سنوات تجاوز الطبيعى وأيضا فى الـ 13 ، نتيجة الطفل فى الأخبار يمكن أن تعتبر عند المصطلح "العمر العقلى" فى الترجمات المتعددة لميزات الاختبار يمكن ان نعبر عن المصطلح "العمر العقلى" استبدل بـ "المستوى العقلى" لأن العمر العقلى يكون مفهوم بسيط للإدراك ، تقديم هذا المصطلح بلا شك فعل للكثير لثمرة

الاختبارات الخاصة بالذكاء "بنيت" ذات نفسه تجنب مصطلح "العمر العقلى" بسبب تطبيقاته التطورية الغير مثبتة وفضل المصطلح المحايد "المستوى العقلى".

النسخة الثالثة لميزان "بنية . سيمون" ظهرت فى عام 1911 ، السنة التى مات فيها "بنيت" فى هذا الميزان ، لم تقدم اى تعبيرات أساسية مراجعات صغيرة وإعادة وضع اختبارات قد استبدلت ، اعتبارات آخرى أضيفت على مستويات عديدة واسع ليشمل مستوى البالغين ، قبل النسخة 1908 ، اختبارات "بنية سيمون" جذبت انتباه واسع بين 4 علماء النفس عبر العالم. الترجمات والاعدادات ظهرت فيمدن كثيرة تضم المعاقين.

اختبار المجموعات:

اختبارات "بنية" فى جميع النسخ هى مقاييس فردية ، وليس لإدراة المجموعة وتحتاج لمتحنين على مستوى عالى ، وهى أساسا وسيلة طيبة تتناسب مع الدراسة المكثفة لحالة فردية.اختبار المجموعات بشبة ميزان "بينت" فى انه قد تطور ليستخدم فى حاجة عملية ، عندما دخلت أمريكا الحرب العالمية الأولى 1917 عينت بواسطة الاتحاد الأمريكى لعلماء النفس لبحث طرق يستطيع فيه علم النفس أن يساعد فى الحرب هذه اللجنة تحت إشراف المستوى العقلى العام ، فى هذا الاختبار قام علماء النفس العسكريون بالاقتراب من كل مواد الاختبار المتاحة وخاصة اختبار جماعى غير منشور أعد بواسطة سوتس "وحولوة إلى الحرب المساهمة الكبيرة وهو تقديم اختبارات متعددة هذه الاختبارات تطورت بواسطة علماء النفس العسكريون وعرفت بالفاو الحربية ، الشكل صمم للاختبارات المعتادة العامة الأخيرة كان ميزان غير لغوى وظلف للاميين والجنود الغير وطنيين وكان الاختبارات مناسبان لمجموعات كبيرة ، بعد الحرب ، الاختبارات الحربية استخدمت مدينا ، لم ليس فقط اختبارات الفاوبيتا مرة بمراجعات عديدة ولكن

استخدمت كنموذج للاختبار الجماعى ، حركة الاختبار تطورت ونمت بشكل كبير قريبا اخبتارات ذكاء جماعية صممت لكل الأعمار والأشخاص من أطفال ما قبل المدرسة إلى طلاب الجامعات والخريجين ، برامج مقاييس الاختبارات الكبيرة التى كانت متاحة من قبل قدمت بشكل كبير أكثر تفاؤلا ، لأن اختبارات المجموعات صممت كوسيلة اختيار ضخمة ، تضم تعليمات مبسطة وإجراءات إدارية تتطلب أدنى قدر من التدريب من جانب الممتحن ، بدأ مدرس المدارس بإعطاء اختبارات الذكاء لفصولهم ، طلاب الجماعات يخيروا قبل القبول ، دراسات مكثفة على مجموعات من البالغين كلما أجريت. تطبيقا تلك الاختبارات عجل بتطوير أساليبها . حيث اختفت الاختبارات الغير جيدة . مع الأعمار بجمع الأهداف والنتائج.

النشأة التاريخية للاختبارات الحديثة:

للتأكد ، الاختبارات تغطى القدرات التى تكون ذات أهمية أساسية فى الثقافة والتى صممت من أجلها ولكن تم إدراك أن التصميمات أكثر دقة ، بشرط نوع المعلومات التى يمكن أن تنتجها تلك الاختبارات ، يمكن أن تكون مفضلة على سبيل المثال ، الاختبارات التى سميت باختبارات الذكاء خلال العشرينيات أصبحت تعرف باختبارات الاستعداد الدراسى. هذا التحول فى المصطلحات حدث مع التعرف على حقيقة أى مسمى "اختبارات الذكاء" تقيس الدمج بين القدرات المطلوبة والمطلوبة فى العمل الأكاديمى.

حتى قبل الحرب العالمية الأولى ، علماء النفس بداو يدركونا لإحتياج إلى اختبارات الاستعداد لتكملة اختبارات الذكاء.

اختبارات "الاستعداد الخاص" تطورت لتستخدم فى مجال الإرشاد المهنى واختيار وتصنيف الموظفون الصناعيون والعسكريون ضمن الاختبارات الأكثر استخداما توجد الاختبارات الميكانيكية والموسيقية ، الكتابية والفنية.

الملاحظة النقدية لاختبارات الذكاء والتى اتبعت توسعها واستخدامها خلال العشرينيات أو حتى إلى حقائق جديرة بالملاحظة.

أداء الفرد فى الأجزاء المختلفة من الاختبارات عادة يؤدى إلى تباين، ويكون هذا واضحا فى اختبار المجموعات والتى فيها البنود وتكون منفصلة إلى اختبارات جزئية للمحتوى المتجانس. مثال الفرد بما يحقق درجة عالية فى الاختبارات الجزئية اللفظية بما يحقق درجة منخفضة فى الاختبارات الجزئية الحسابية إلى حد ما، مثلها والذى فيه كل Standferd Binet الاختبار يظهر فى اختبار مثل البنود تتضمن كلمات ربما تمثل صعوبة لفرد بعينة.

مستخدمى الاختبارات وخاصة الأطباء يستخدمه للحصول على فهم عميق لنفسية الفرد. وهكذا ليس فقط معامل الذكاء ولكن الأداء فى مجموعة من البنود تختبر فى ملاحظة حالة الفرد. مثل التمرين لا ينفصل أى يستخدم بشكل عام لأن اختبارات الذكاء تصمم من أجل تحليل الاستعداد. عادة الاختبارات الجزئية تتضمن بنود قليلة لإنتاج قدرة خاصة مستقرة. كنتيجة الإختلاف بين أهداف اختبارات الذكاء ربما تنعكس لو أعيد اختبار الفرد فى يوم مختلف أو بشكل آخر من نفس الاختبار.

بينما توضح الاستعمال للاختبارات الحاجة إلى اختبارات استعداد متعددة، التطور الموازى فى دراسة تنظيم السمات كان يعطى الوسائل تدريجيا لبناء الاختبارات. الدراسات الإحصائية على طبيعة اكتشاف العلاقات الداخلية بين الأهداف التى تحصل عليها عن طريق تنوع اختبارات الأفراد مثل هذه البحوث بدأت عن طريق العالم النفسانى الإنجليزى (شارلس سبيرمان) خلال العقد الأول من القرن العشرين. التطورات فى علم الطرق بينت على أعمال علماء النفس الأمريكان مثل (تل ككلى، ترستون) وأيضا على علماء آخرون إنجليز وامريكان عرفوا باسم (فاكتور إنليس) التحليل العاملى.

المساهمات التى قدمتها طرق التحليل فى بناء طرق التحليل العاملى فى بناء الاختبار سوف نوضحها فى الفصل الحادى عشر اما الآن سوف نلاحظ أن المعلومات التى جمعت بمثل هذه الطرق أشارت إلى وجود عدد من العوامل المستقلة أو المميزات. بعض هذه الميزات (السمات) تتمثل فى اختبارات الذكاء التقليدية ، المحادثات اللفظية والتفكير العددى تكون أمثلة لهذا النوع من السمات.

واحدة من العائدات العملية الرئيسية من "التحليل معاملى" كان بطاريات الاستعداد المتعددة صممت هذه البطاريات لتقديم مقياس لمستوى الفرد فى كل سمة.

عوضا عن النتيجة الكلية أو معامل الذكاء ، نحصل على نتيجة منفصلة لمثل هذه السمات مثل المحادثة اللفظية ، السلوك العددة حيز الرؤية ، مثل هذه البطاريات تعطى أداة مناسبة لعمل تحليل فردى أو تشخيص تباينى (تفاضل) والتى حاول مستخدمى الاختبارات يحصل عليه لسنتين طويلة من اختبارات الذكاء هذه البطاريات أيضا تدمج فى برامج الاختبارات المنظمة.

البطاريات للسلوك المتعددة تمثل تطورا نسبيا فى مال الاختبارات ، قريبا ظهرت منذ 1945 وفى هذا الاتصال عمل الأطباء النفسيين العسكريين خلال الحرب العالمية الثانية يجب أن يلاحظ كثير من البحث فى الاختبارات وجه لخدمة الجيش وتلك مبنيا على التحليل العاملى ووجه إلى بناء بطاريات الاستعداد المتعددة فى القوات الجوية ، بطاريات خاصة صممت للطيارين وكثير من المتخصصين العسكريين.

تطور أكثر حداثة ، ظهر مؤخرا فى أواخر الثمانيات وبداية التسعينات تعطى تكامل أساسى لطريقتين سابقتين متعارضين للقياس الذهنى اللذان يمثلان فى اختبارات الذكاء التقليدية وبطاريات الاستعداد المتعددة سوف يأتى

ليدرك أن القدرة الإنسانية يمكن أن تقيم على مستويات مختلفة ، من التعرف الضيق للاستعداد فى اختبارات محددة إلى نتيجة نهائية مثل معامل الذكاء التقليدى ، من الأغراض المتعددة للاختبار المتطورة مثل موازين القدرة المتباينة أو المراجعات الحديثة للاختبارات الأولى مثل الصوقر الرابعة لستانفورد تمتزج بتغطية الاستيعاب للاستعدادات المختلفة. وعلى الرغم من أن كلا من المثالين السابقين اختبارات ذكاء للفرد إلا أنها لها تأثير على المجموعات.

اختبارات الإنجاز الموحدة:

بينما كان العلماء النفسيون منشغلون فى تطوير اختبارات الذكاء واختبارات الاستعداد وكانت اختبارات المدرسة التقليدية تمر بعدد من التطورات كخطوة هامة فى هذا الاتجاه أخذها مدرسة بوسطن عندما استبدل الاختبار المنطوق باختبار مكتوب للتلاميذ. من المناقشات التى تدعم هذا الابتكار أن الاختبار المكتوب يضع التلاميذ فى موقف واحد ويسمح بتغطية جميع المحتوى ، يقلل من عنصر المجازفة ، فى اختيار الأسئلة والتخلص من إمكانية التحيز للمتحن.

بعد مرور القرن ، أول الاختبارات الموحدة لقياس عائدات التعلم المدرسى بدأت فى الظهور تلك الاختبارات ، استخدمت مبادئ القياس المتطورة فى مختبرات علم النفس.

مثال على ذلك موازين مقياس لتصنيف جودة الكتابة واختبارات الإملاء والتفكير الرياضى. مؤخرا بطاريات الإنجاز بدأت مع الانتشار النسخة الأولى من اختبارات ستيفورد للإنجاز 1923 واصحابها كانوا ثلاثة رواد فى تطور الاختبارات.

وكانت بداية العديد من الخصائص للاختبار الحديث ، هذه البطاريات تعطى مقاييس مشابهة للأداء فى مواد المدرسة المختلفة وفى نفس الوقت كان هناك دليل على نقص الاتفاق بين المدرسين على تصنيف الاختبارات المقالية.

عام 1930 أدرك أن اختبارات المقال ليست فقط مستهلكة للوقت ولكنها أيضا تنتج نتائج لا يمكن الاعتماد عليها.

كلما مر الوقت زاد استخدام اختبارات الإنجاز الموحدة وكان هناك تأكيد متزايد على تصميم البنود لاختبار الفهم وتطبيق المعرفة والأغراض التعليمية الآخرى.

العقد الثالث من القرن العشرين شاهد تقديم واستقرار برامج الاختبار وكان تطورا جديرا بالملاحظة ربما كان الأشهر من هذه البرامج هو محليا ودوليا الذى كان نقطة تحول لتقليل تكرار اختبار المبتدئين ، هذا البرنامج شاهد تغيرات فى اجراءاته وفى عدد وطبيعة فى السنتين Ceeb المشتركين ، فى عام 1947 وظائف اختبار اللاحقة تحملت مسئولية ازدياد عدد برامج الاختبار بالنيابة عن الجامعات ، المدارس المحترفة والوكالات الحكومية.

اختبارات الإنجاز استخدم فقط للأغراض التعليمية ولكن أيضا لاختبار طالب الوظائف الحكومية والصناعية نذكر مثلا الاستخدام المنظم فى الصين لاستخدام اختبارات الخدمة المدنية.

فى الوقت الحديث ، اختيار موظفى الحكومة بالاختيار قدمت فى أوروبا فى أواخر القرن 18 وأوائل القرن الـ 19.

مزيد من العلماء شاركوا فى بناء اختبارات الانجاز الموحدة ، السمات الفنية لاختبارات الانجاز التى تمثل اختبارات الذكاء واختبارات الاستعداد.

الإجراءات لبناء وتقييم كل هذه الاختبارات أصبحت متداولة جدا والاهتمام المتزايد لأعداد اختبارات الإنجاز والتى يمكن أن تقيس تحقيق أهداف تعليمية واسعة جعل محتوى اختبارات الإنجاز تشابه إلى حد كبير اختبارات الذكاء. اليوم الاختلاف بين هذين النوعين من الاختبارات يكون رئيسا واحدا من الدرجة التى يكون عليها المحتوى من التخصص إلى أى مدى الاختبار يفترض مقدما التعليمات التى لها أولوية.

تقييم الشخصية:

مساحة أخرى من الاختبار النفسى تعتنى بالمظهر المؤثر للسلوك الغير فكرى ، وهذا سوف يناقش فى الفصل 13 - 16 صممت اختبارات لهذا الغرض وتعرف باختبارات الشخصية على الرغم من أن علماء النفس يفضلوا أن يستخدموا مصطلح "الشخصية" بمفهومها الواسع للإشارة إلى كل ما هو داخل الفرد.

السمات الفكرية والغير فكرية تدخل تحت هذا العنوان فى مصطلحات اختبارات علم النفس ، تصميم "اختبار الشخصية" يشير عادة إلى مقاييس لمثل هذه الصفات كالحالة العاطفية ، العلاقات الشخصية ، الدوافع ، الاهتمامات والسلوكيات والاتجاهات.

بادرة لاختبار الشخصية يكون واضحا باستخدام اختبار الاتحاد الحر مع المرضى النفسيين. فى هذا الاختبار ، الممتحن يعطى كلمات ويطلب منه الإجابة على كل منها بأول كلمة تأتى على باله.

(كى بلين) وظف هذا الأسلوب ليدرس التأثير النفسى للإرهاق ، الجوع والأدوية "المخدرات" وختم بأن كل هذه العوامل تزيد التكرار النسبى للاتحاد الظاهرى.

1894 (سومر) اقترح أن اختبار الاتحاد الحر ربما يستخدم ليفرق بين لأشكال المتعددة للتشويش الذهنى.

أسلوب الاتحاد الحر وحد من أجل تنوع إغراض الاختبار ومازال موظفا نذكر أعمال كاتيل ، برسون ، جالتون فى تطور الاستفسار الموحد وموازين التصنيف. وعلى الرغم من أنها أساسا ابتكرت من أجل أغراض آخرى إلا أن تلك الإجراءات وظفت عن طريق آخرون فى بناء الأنواع الشائعة للاختبار الشخصية الحديثة.

النموذج الأول لاستيفاء الشخصية ، هو "ورقة عمل المعلومات الشخصية" المطورة عن طريق "ودورث" خلال الحرب العالمية الأولى.

هذا الاختبار صمم للتعرف على الرجل المضطربين عقليا الذين لا يصلحون للخدمة العسكرية. هذا البيان يتكون من عدد من الأسئلة التى تتعامل مع الأغراض الشائعة للأمراض النفسية والذى يجيب عنه أنفسهم. النتيجة النهائية نحصل عليها بعد الأعراض المعطاه. تطور "ورقة عمل المعلومات الشخصية" لم يستكمل بشكل كاف ليسمح بالاستخدام العملى لها قبل نهاية الحرب مباشرة ، أشكال مدنية أعدت لتتضمن شكل خاص يستخدم الأطفال.

فى بعض الاستفتاءات ، محاولة لتقيم التكيف العاطفى فى أشكال أكثر تخصصا مثل التكيف المنزلى ، التكيف المدرسى ، التكيف المهنى ، اختبارات آخرى ركزت بشكل أكثر حدة على مساحة ضيقة للسلوك أو اعتنت بالاستجابات الاجتماعية ، تطور آخر كان بناء اختبارات قياس تغيرات السلوك والاجتماعات ، تلك الاختبارات أيضا بينت أساسا على أساليب الاستفتاء. الممتحن يأخذ عمل ليؤديه ويكون الغرض منه غير معروف عادة ، أول تطبيق شامل لهذه الأساليب يوجد فى الاختبارات المطورة فى أواخر العشرينات وبداية الثلاثينات عن طريق

(هارثورن) هذه السلسلة وحدت على أطفال المدارس وأعتنت بسلوكيات مثل الغش والكذب والسرقة والتعاون ، توضيح آخر على مستوى البالغين أعطى عن طريق سلسلة من الاختبارات الموق فيه تطورت خلال الحرب العالمية الثانية فى برنامج التقييم لمكتب الخدمات الاستراتيجية.

تلك الاختبارات اعتنت بالسلوك الاجتماعى المعقد العاطفى وتطلبت تسهيلات وأشخاص مدربين لادراتها ، تفسر الاستجابات للفرد كان متصلا بهذا الموضوع.

الفصل السادس

نظرية الاختبار

أولا: ما هى نظرية الاختبار

أينما كان لتصبح فى صميم الموضوع لابد من معرفة صريفية تحتوى على تعزيز كمى أحسن وأجود (ثورنديك).

مبدئيا الدارس الذى يتطرق لدراسة نظرية الاختبار غالبا ما يواجهه أسئلة منها:

1- ما هى نظرية الاختبار وفيما نحتاجها.

2- ما هو وجه الاختلاف بين تقييم التعلم واستخدام توعية الاختبار.

3- هل ظرية الاختبار ذات منشئ نفسى أم طريقة جديدة للاختبار.

4- كيف تكون نظرية الاختبار وثيقة الصله بتصميم البحث والإحصاء ، هذا الجزء سيتضمن إجابة الأسئلة فى نهايته وسيتبع إيجاد النقاط الأساسية لوصف المشكلات الخاصة التى يواجهها القياس بشرح جزء كبير من أغراض نظرية الاختبار وعلاقته بهذه المشاكل واستعراض المنشئ التاريخى لهذا الرع وأخيرا سوف نحدد الفرق بين نظرية الاختيار والأوجه الأخرى العلوم الاجتماعية وتقيمها.

تعريفات مختلفة للقياس:

لفهم فائدة نظرية الاختبار لابد أولا من فهم بعض الأشياء عن الجوهر الأساسى لتركيبه القياس والاختبارات النفسية ، وهناك تعريفات متعددة للقياس منها:

1- وايترن هوفر 1951 : يشرح القياس كعملية تتمثل فى العالم الطبيعى عن طريق الملاحظة.

2- ستيفن 1946 : القياس عدد من المهام الهادفة تصاغ وفقا لقوانين محددة.

3- لورد ونوفيك 1958 وتيرجستون: أكثر دقة وإحكام من تعريف ستيفن عن طريق القياس القادر على تسوية تهدف لحل التعارض مع الآخرين ومع نفسه فمثلا فى الكيمياء الطبيعة الكيميائية تقيس مقدار التركيب الجزئى (الحجم. الكتلة) فى الأحياء تحدد المقاييس التى تحدد الوسط الجرثومى عند نزول بركة ماء مثلا أى كل مجال يحدد نوع القياس الذى يميزه.

كيف ظهرت الحاجة للقياس النفسى:

جميع العلوم تسعى أساليب موضوعية دقيقة لقياس الظواهر المتعلقه بموضوع دراستها من أجل فهم الظواهر وتفسيرها والتنبؤ بالعلاقات القائمة بين متغيراتها ومحاولة ضبطها والتحكم فيها لذلك تعددت التعريفات بتطور الدراسات والتطور القياسى كانت التعريفات التالية:

أولا: المفهوم العلمى للمقياس:

1- هو الجوانب الكمية التى تصف خاصية أو سمة معينة مثل (ارتفاع/حجم/استعداد لفظى لدى طفل/تحصيل دراسى لدى طالب) يشير إلى عملية جمع المعلومات وتنظيمها بطريقة مرتبة بذلك يتضمن مفهوم القياس من هذا المنظور كلا من عملية جمع المعلومات وتنظيمها وكذلك نتيجة هذه العمليات.

2- تعريف آخر: تعيين فئة من الأرقام والرموز متأخرة لفئة من الخصائص أو سمات الأفراد طبقا لقواعد محددة تحديدا جيدا حيث لا نستطيع قياس الأفراد ذاتهم.

3- تعريف آخر: هو الوصف الكمى الموضوعى للأداء فهو الحكم الكيفى الوصفى على الدرجة مماثلا فى التقدير النوعى للأداء وهذا الحكم يفيد فى إتخاذ قرار معين بشأن الفرد الذى حصل على الدرجة و اقتراح إجراء مناسب له.

فهذه عملية منهجية تتطلب جمع بيانات موضوعية ومعلومات صادقة من مصادر متعددة باستخدام أدوات القياس المتنوعة فى ضوء أهداف محددة بغرض التوصل إلى تقديرات كمية على أساسها يتخذ قراراته.

4- تعريف أنستازى: هو مقياس موضوعى ومقنن لعينة من السلوك والاختبارات النفسية ليست إلا عينة صغيرة ممثلة للسلوك القياس النفسى والتربوى (د/محمد عبدالسلام).

ثانيا: أغراض القياس النفسى:

المسح . التنبؤ . التشخيص . العلاج.

ثالثا: تطبيقات القياس النفسى والتربوى:

1- التربية والتعليم (التأخر الدراسى . الاختبار التربوى . التوجيه التربوى . الإرشاد التربوى . برامج الاختبار. القراءة العلاجية . تصنيف التلاميذ).

2- فى الصناعة: (اختيار العمال . التقييم . التدريب . الحوادث والتأهيل المهنى).

3- فى الجيش: على اثر الحرب العالمية بدء الاتجاه نحو القياس لاختيار الجنود التى تانسب مواقع معينة وذلك يتم من خلال تطبيق المقاييس النفسية.

4- فى علم النفس العيادى.

5- فى الإدارة (القياس النفسى والتربوى د/محمد عبدالسلام).

السمات النفسية

تعريفها:

حاول علماء النفس اختزال التعريفات فى إطار ما يعرف بنظرية السمات والسمة مفهوم يصف السلوك.

1- السمات عبارة عن:

تجمع من السلوك المترابط (أى حدوثة معا).

السمة: مفهوم مجرد ليس شيئا ملموسا فهى عامة (مثل الميل للمشاركة فى المواقف الاجتماعية) فهى لا يمكن قياسها مباشرة وإنما يستدل عليها من انماط السلوك الملاحظ.

2- السمات العقلية:

تعد بمثابة خصائص يفترض إنها تعكس مجموعة مترابطة من السلوك الذى يمكن ملاحظته وتسجيله فى مواقف اختبارية مقننة يجب أن تفى السمات ببعض الشروط لكى تتم الإفادة منها فى وصف السلوك.

شروط السمات التي تصف السلوك:

1- أن تمثل السمة خاصية محددة يتباين فيها الأفراد مثل السمات العقلية الشخصية والوجدانية.

2- أن تكون السمة قابله للتعريف الإجرائى بحيث يكون هناك قدر معقول من الإتقان بين القائمين بملاحظة السلوك حول مقدار السمه لدى الأفراد

المختلفين فمثلا ما الذى نعنيه بسمة الذكاء ما نوع السلوك الذى يميز الشخص الذكى ويسمه؟

3- أن تتميز السمة بالثبات النسبى بمرور الزمن باختلاف المواقف فسلوك الفرد يصعب التنبؤ به من موقف إلى آخر.

الخصائص الأساسية للسمات:

حدد تورنديك ونانلى مجموعة من الخصائص الأساسية للسمات التى تقيس عمليات قياس السلوك الذى يستدل بها على سمة معينة وفيما يلى هذه الخصائص الأساسية للسمات التى تسير عمليات قياس السلوك الذى يستدل منه على سمة معينة وفيما يلى هذه الخصائص:

1- يمكن أن تمثل السمة بخط متصل: يمكن تتبع سمة الذكاء فى مستويات مختلفة متدرجة يتمايز فيها الأفراد بحسب درجاتهم فى هذه السمة.

2- يمكن تدريج متصل للسمة بوحدات مناسبة مثل (عبقرى. ذكى جدا. فوق متوسط. متوسط الذكاء. دون المتوسط).

3- يمكن قصور متصل للسمات على إنه ميزان قياس فترى.

أنواع السمات:

(سمات تتعلق بالشكل الخارجى للجسم . سمات فسيولوجية . سمات مزاجية9 الاستعدادت استعداد (دراسى . مكانى . ميكانيكى .رياضى . موسيقى . فنى) مهارات التحصيل. ميول. اتجاهات. قيم.

علم النفس التقليدى:

1- يحدد كما يدرس السلوك الفردى للشخصية (فى البيت . فى العمل . فى المدرسة . أو المواقف الاجتماعية العامة) تسمى سمة نفسية.

2- فالسمة النفسية غير متشابهة فهى فردية لا نستطيع قياسها مباشرة.

كأن تكون (عالى ، منخفض . خفيف أم ثقيل) فهى مركبة وقبول هذه الفرضية ناتج عن نماذج من علم التخيل فى العلوم الاجتماعية التى تحتوى على تطوير نظريات دراسة السلوك الإنسانى إن تواجد فى المجتمع.

وكينونة وجود مثل هذا المنشئ لا يستطيع أبدا أن يجزم أو يؤكد على ذلك فدرجة البناء النفسى الفردى للشخصية نستطيع أن نستدل عليها فقط عن طريق ملاحظة السلوك دعنا نحدد ذلك من خلال مثال فعملية تكوين المعلومات مثلا وكيف تؤدى إلى قياس الخصائص قياس السمات النفسية تستلزم تطوير ضمنى للعمل فى ملاحظة تنوع أنشطة لعب الأطفال فى الروضة أثناء محاولتهم تقليد الآخرين مثلا.

فبعد ملاحظة هذا النوع من السلوك (تقليد الكبار) غالبا ما يكون متناسق فى البيئات المختلفة المتماثله فرديا.

3- تصنيف السلوك كبعد اجتماعى: فالتعريف النفسى يحتوى على تركيب بناء (نظرى) يقوم بترقيم (تحويل السلوك الأرقام) لعينة من السلوك لكن محتوى ذلك البناء قد لا يماثل القياس فى بيانه لذلك.

تركيب بناء المقياس

قبل إعداد أى مقياس يجب أن تكون لدينا قدرة على إيجاد القواعد الأساسية فى التوفيق بين البناء النظرى وملاحظة السلوكيات ليكون منطقى

معبر عن التركيب وهذه العملية تسمى إيجاد تأسيس فعال فعلى سبيل المثال لقياس البعد الاجتماعى يجب على الأخصائى النفسى إتباع ما يلى:

1- ان يحدد ما أنواع السلوك أثناء اللعب فى الحضانة وما يحتويه من أبعاد.

2- يقترح توصيات (خطة) للحصول على عينة سلوكية فى مستوى المواقف الفعلى.

3- مراقبة سلوكيات كل الأطفال فى مستوى نموذجى فذلك يتطلب وسيلة متطورة (أو وسيلة فى بناء البعد الاجتماعى).

4- هذه الوسيلة توجد مستوى إجرائى منهجى للحصول على عينة سلوكية من خاصية البعد وكاستخدام لهذه الجزئية فى هذا النص يرجع الاختبار إلى إجراءات للحصول على افضل عينة نموذجية فردية. كإشارة للاتجاهات أو اقسام الاختبار التى تختبر البناء الأفضل أو العينة الفردية النموذجية المتطابقة.

القياس النفسى خاصية تحدث عندما:

عندما يتم التقييم الكمى لسمات عينة ا لسلوكيات التى جمعت عن طريق استخدام الاختبارات أى تحويل السمات النفسية لأرقام وكذلك الأبعاد الاجتماعية.

لماذا الاخصائيين النفسيين والمعلمين يؤكدون أن بناء القياس النفسى لا يمكن أن يحدد من خلال الملاحظة المباشرة؟

ففى التعليم والمجال النفسى يشترط البناء الفعال والطرق الملائمة لتصنيف أرقام السلوك المتشابهة وبدون مثل هذا البناء عملية تصنيف ووصف السلوك الجزئى غالبا ما تحدد بملاحظة معقدة للسلوك تظهره بطريقة تدريجية تشويش وإرباك فلشرح عدد كبير من الكلمات والأفعال وردود الأفعال يمكن أن

يوجد فى يوم منفرد (توفير يوم مستقل وفصل مستقل بكلا من المعلم والأخصائى ليقوم بملاحظة دقيقة وشرح وتفسير سلوكيات التلاميذ).

فمن خلال استخدام البناء حينما تكون الملاحظة قادرة على تصنيف وتجميع أمثله متشابهة من السلوك وتوصيل الأجزاء المدمجة المشاهدة مثل هذه الأبنية كحجر أساس لبناء نظرية عن السلوك البشرى.

فعلى أبعد مستويات نظريات علم النفس التي تعبر استحاله وجود علاقة بين اثنين من تركيبات البناء النفسى أو بين البناء والملاحظة الظاهرة فنتائجها قابله للتطبيق وأخيرا من خلال التدارس التجريبى ثم إثبات مثل هذه النظريات التى تؤدى لاحتمالية التنبؤ والتحكم فى نماذج سلوكية معينة لإنجاز هذه الأهداف ومع ذلك من الضرورى دوما التيقن من ملاحظة السلوك الذى يشرح بناء الفروض فى هذه النظرية.

ثالثا: مشاكل بناء القياس النفسى

لأن البناء النفسى هو التجريد الذى يستطيع فقط تقييم صعوبة التصميم فى معدات القياس فلابد من توضيح بعض أسباب الاحتياج للقياس النفسى.

1- الأشخاص العاملين بالمجال النفسى يريدون تطوير ميكانيكى يتجه نحو مقابلات الوظيفة فى رواسخ مشتركة.

2- المدارس النفسية ترغب فى تطوير مهارة التخمين لدى المعلم فذلك يحتاج لدراسة نفسية معقدة.

3- 50٪ من المدرسين يرغبوا فى تطوير وحدات اختبار التلاميذ لقياس مهارات أطول وأكثر على أوسع نطاق ن خاصتا الاتجاه الميكانيكى حيث يكون الأداء يدوى معقد فجزء كبير من هذه المهارة يختلف تماما عن حالة الاخرى.

التطوير فى الاختبارات يواجه غالبا 5 انواع من المعوقات او المشاكل المتعلقة بالقياس:

1- لا يوجد طريقة فردية موحدة لفهم بناء القياس بشكل عام ومقبول.

2- غالبا ما يحدد بناء القياس النفسى عينه من السلوك.

3- المقاييس غالبا ما تحرز أخطاء (إستجابات خاطئة).

4- يوجد قصور فى إيجاد جزئية المقاييس المتدرجة فذلك يسبب مشكلات من نوع آخر بناء القياس وحده لا يستطيع حلها.

5- صعوبة تحديد العلاقة بين الأبنية المختلفة للمقياس والملاحظة الظاهرة.

أولا: لا يوجد طريقة عامة موحدة لفهم بناء القياس النفسى:

لأن قياس التراكيب النفسية دائما ما يكون قائم على أساس السلوك بشكل غير مباشر ولذلك لابد من فهم وثيق الصلة بالبناء المطروح دراسته فمن المحتمل دائما ثنائية الاختبارات أكثر من تماثلها حيث يتم اختيار عناوين مختلفة من السلوك الموجودة فى الأبنية (المستخدمة) الجاهزة للعمل. فعندما نتحدث عن تطوير الاختبارات التى نرغب من خلالها تقييم مهارة التلاميذ على المدى الطويل بعد ذلك لم يستطيع أحد أن ينظر مباشرة لما فى داخل رؤوس التلاميذ وينظر كم يبلغ طول مدى معرفتهم ، فالاختبارات المتطورة يجب أن تصمم ؑلى تصنيف السلوكيات الظاهرة للطلبة القادرة على إظهار طريقة الاستدلال عن مدى معلوماتهم فى المواد المختلفة.

ثانيا: القياس النفسى غالبا يحدده عينة من السلوك:

فعلى سبيل المثال يتعذر التعامل مع اختيار (مقارنة) مواجهة التلاميذ للمشاكل المحتملة على المدى الطويل لأن ذلك يمكنهم من توقع الحلول لذلك

تحديد أى قياس. لمهارتهم فى هذا النطاق يجب أن يتضمن مثل هذه المشكلات البسيطة فتذكر أرقام الوحدات وتنوع المحتوى شرط ضرورى لأبسط تعلم أبعاد السلوك فتكون مشكلة تطور القياس هى معنى إجراءات القياس.

ثالثا: المقاييس دائما ما تحرز استجابات خاطئة:

معظم المقاييس النفسية ترتكز على تحديد الملاحظة البسيطة ودائما تتناول نقطة واحدة فقط أو حدة السلوك فعندما نتأمل مرة أخرى التلاميذ الذين أخذوا اختيارات على المدى الطويل إذا أخذ نجاحهم فى اختيارين متماثلين فعدم التشابة فى درجاتهم سيكون منطقى ذلك يرجع لتأثير (التعب . الملل . النسيان . الإهمال...) فمشكلة أخطاء إجراءت القياس أساسية فى القياس النفسى.

رابعا: يوجد نقص فى المقاييس المتدرجة:

يوجد نقص فى إيجاد وحدات قياس متدرجة بشكل مشكلة آخرى (فى الحقيقة هذه الاختيارات قد تستطيع إجابة أسماء هذه الأجزاء على المدى الطويل إلا أن متطابق نتائجها = صفر) فهل عامل الاختيارات يستطيع الإجابة إجابة مطلقة لهذه الفقرات على المدى الطويل مثلا إذا جاب طالب يدعى سيو على 5 فقرات وجوى على 10 فقرات صحيحة وستيف اجاب 15 فقرة فهل تستطيع إفتراض إختلاف الكفاءة بين سيوه 5 وجوى 10 درجات وهل يشبه الاختلاف بين جوى وستيف.

خامسا: بناء القياس لا يستطيع إيجاد حلول لهذه المشكلات:

لكى أيضا يجب أن يحدد العلاقة بين التركيبات الاخرى والملاحظة الظاهرة فالقياس النفسى يحدث من خلال الاعتماد على الملاحظة الشخصية مما يقلل من المعانى وفائدتها فلا تستطيع تتبع حقيقة ما يتضمنه بناء النظرية.

ثالثا: خطة نظرية الاختبار:

دراسة المشاكل المتعددة فى القياس تكون وصفية فقط (غالبا) فلابد من إيجاد حلول لهذه المشاكل التى تستخرج من تخطيط خاص من الجانب التعليمى والنفسى وذلك للتعرف على نظرية الاختيار. ويحتوى هذا التدريب على قواعد أولية لهذه الطرق منها:

1- الحد من تأثير هذه المشاكل على جودة القياس.

2- من المفترض فى دراسة نظرية الاختيار ان يساعد دراسية ليصبحوا على وعى منطقى بنماذج رياضية تتدرج تحت مستوى جيد من استخدام الاختبارت وبنائها.

3- بعض الطرق التجريبية والمقاييس الاحصائية تستخدم فى التعليم والأبحاث النفسية التى يكون منشئها الأبحاث.

4- عند ذلك يتحتم تقييم تأثير المعالجة التجريبية. تقييم كمى وكيفى.

رابعا: التطور التاريخى للقياس النفسى:

التطور التاريخى للإختيارات يمزج بين التطور فى علم النفس لخاصية الاختبار حيث إتجه التطور النفسى فى أوربا وأمريكا لدراسة متنوعة لعلم النفس ومشكلات التعلم ففى هذا الجزء سيتم توضيح بعض التركيبات القليلة ، كمثال فردى توضيحى لرجال وسيدات يعملون فى مجال نظرية الاختيار والمقابلات.

1- فى المانيا:

بدء فى منتصف الثمنيانينات 1800 فى معمل الإدراك الحسى فى ليبزخ فى المانيا مكثفها (ويلم وندت. إرنست ويبر . جيسنق فيشنز) هذه المجموعة شرحت

اهم الترتيبات الأولى على قياس نفسى مع الإهتمام بالتحكم والضبط التحكم والضبط الإجتماعى. وأن الدراسات النفسية الأولية اتخذت من إجتهادات فلسفية وتعريفات عامة فى وصف التقييم الطبعى فى هذه العملية.

2- بريطانيا:

إهتموا أكثر بدراسة الفروق الفردية وأول من وضع ذلك فرنسيس جالتون حيث أفكاره كانت ذات تأثير قوى فى عصره وتشابه مع نظرية دارون للوراثة.

جالتون:

إهتمامه كان متنوع وبعض طرقه تبدو غير تقليدية فمثلا (مستوى اليوم) فى إحد مقالاته وصف نماذج عامة تحتوى اهداف تستطيع أن تنسيئ فيما بعد عن السلوكيات المتوقعة للأفراد من خلال خصائصهم الوراثية.

وأيضا أسهم فى تسجيل تساؤلات متقدمة فى القياس النفسى عند دراسته الشخصية. سنة 1869 بإستخدام الدرجات من إمتحانات الرياضيات فى الكلية الحربية (القوات الملكية) ثم التوصل إلى أن ابعاد القدرة القلية من المحتمل أن تقدر بتوزيع تقريبى (توزيع تقدير تقريبى) على المنحنى الإعتدالى. عموما فى العشرين سنة الأخيرة ثم تقنين طبيعة الفروق الفردية حيث وضح جالتون إقتراحات سهلة لإجراءات متبادلة لإختبارات لفظية أكثر مساحة.

كارل بيرسون:

طور فيما بعد ذلك نماذج إحصائية وثيقة الإرتباط بهذه الإقتراحات (اقترحات جالتون) ويعتمد عليها معظم العلماء البريطانيون فى ملاحظاتهم.

سبيرمان:

هو من صاغ نظرية الذكاء كأساس أكثر تقدما فى الإجراءات المتبادلة يسمى التحليل العاملى وعوامل التحليل تظل وسط إحصائى شائع يستخدم المصدافية الإختبار.

3- فرنسا:

رغم أهمية الإسهامات السابقة للنفسيين الألمان والبريطانيين أنقسم الفرنسيين لجزئين الفريد بنيت وثيوفلى سيمون 1908/1905 التى إتجهت دراسته نحو الإختبارات العقلية من أكاديمية التدريب العسكرية ذلك تستطيع وضع عمل مقابلات فى الفصل أورش العمل والقسم الثانى الأخصائيين النفسين العيادين حديثهم يدور حول مهمة البرامج المصممة وطرقها حيث أنها ستفيد لتعريف خاصية القدرات العقلية للأطفال المدرسة العامة (الداخلية) والنتائج كانت ناجحة فى إجراءات القياس المسحية التى شملت الموضوعات العامة (الذكاء) على الرغم من ذلك ربينت إشتهر كثيرا بالإختبارات التى أنتجت تحت هذا الأسم يؤدى تأثيرا فى نظرية الإختبار فى نطاق العمل البناء فى تقييم الإختبارات. قدم بنيت إستنتاج مقبول وعام لتركيب الإختبار حيث كان ينظم ضبط الوحدات الأساسية فى إطار نظرى منطقى جاهز للاستخدام.

4- امريكا:

على الرغم من أن المدارس الأمريكية كانت نشطة فى إيجاد حلول المشاكل النفسية فى أواخر الثمانينات فلم تكن بدءت بعد فى إيجاد طرق القياس حتى القرن العشرين فعلى سبيل المثال كانت فى محاولاته لصياغة اختبارات القدرات العقلية فى كتابات مبسطة سنة 1890 فى هذه الكتابات كانت تعريفات دورية عامة بأكثر الإختبارات بساطة التى تحتوى على تقدير معيارى واكد على ان علماء النفس يجب أن يدرسوا الأخطاء الواردة فى الملاحظة. سنة

1933 بويونت وكاتل درسوا الإختبارات الوصفية كان فى المقام الأول الخصائص السيكومترية أو الإدراك الحسى حيث تتشابة تماما الطريقة التى استخدموها مع الطريقة التى استخدمها الألمان فى إختبارات القدرات العقلية فى سنة 1904 ثورزديك كان غير واثق من مثل هذه الاختبارات لكنة توصل فيما بعد لمجموعة سمات بشرية من خلال بعض المقاوف التى تعبر عن نموذجه (الإناث) العقلى. ويلم جيمس لدى تدوين ملاحظة (جونكش 1968) انا أرسلت لك قصة التى كتبتها (اكتبها ان) فهى لا تنتهى علميا لكن لا تخلوا من الإشارة للإهتمامات الإنسانية فيجب عليك عمل كل أبحاثك بحيث نفى كل ما قرائته لكن لن تظهر بدون تغطيتك الشخصية فالمؤشرات والمنحنيات والنماذج سوف تقودك للغضب الشديد. ارتكز تورنديك فى قياس المشكلات النفسية على نظام بناء كان له صدى كثير فى الدراسات النفسية قدراسة نظرية الاختبار ستصبح مطلب أساسى لتدريب كثير من الأخصائيين النفسين الصغار (المبتدئين) فى كلومبيا كان هارفرد وسنتافورد مبنه ابحاثهم تشابه أبحاث كاتل وثورفديك حيث استأثر بتخيل كليات صغيرة ووضع الترتيبات المشتركة لتعميم تطبيق النظريات السيكومترية فتطوير هذه الاختبارات إعتمد على العمل المتواصل الخاص بحجم الإنتاج أثناء ذلك إتجهة الدراسة فى المجتمع الأمريكى نحو التقييم الإجرائى وانتجت أول مجموعة من الاختبارات سنة 1971 فى نهاية الحرب العالمية أكدت لجنة الأخصائيين النفسيين (بنيامين . جودارد . هنيسن . تيرمان . ويل . وابيل . ويركس) بتصنيف رجال الجيش وفقا لأقسام معينة. و 8700 من اقسام الحرب عملوا لمدة 3 اشهر لإنتاج 5 نماذج لمجموعات إختبارية التى كانت تحدد كل اشخاص الجيش المجموعة الأخيرة من الاختيارات اللفظية كانت أيضا تفضل الإختيارات لمن لا يستطيعون قراءة الإنجليزية.

كانت تستخدم لإنتقاء وتحديد المكان المناسب للجنود المتماثلين فى أفضل الترشيحات فى التدريب المكتبى ويستبعد ذوى العقول غير اللائقة وتحدد علامات ساحة الحرب.

(يركن 1921) أفراد الجيش الذين إجتازوا هذه الإختبارات بنجاح الأخصائيين النفسية فى أمس الحاجة لنشر إستخدام مثل هذه الإختبارات (إختبارات الذكاء) فى التوجية المهنى وتحديد إتجاهات التعليم ، إلا أنه مجهود متميزه وجهة نقد لمستوى الاختبارات التى استمرت أكثر من خمسين عاما (كارون جاش).

تيرستون كاف (1929) اضاف تقنية آخرى لإنتقاء إتجاهات للتقدم نحو تطوير الإختبارات نيرستون والآخرين مثل كيلى وهملتز نيجر حيث قاموا بتطوير تحليل عاملى جديد أفضل.

خامساً: دور نظرية الإختيار فى الأبحاث والتقييم:

أ- الإختيار:

هو أداة قياس مقننة وأسلوب منظم يصمم إضافات للحصول على قياس موضوعى لعينة من السلوك بهدف موازنة أداء الفرد بمعيار أو بمستوى أداء محدد يشمل هذا التعريف.

ب- القياس:

هو الوصف الكلى الموضوعى لأداء فهو الحكم الكيفى الوصفى عالى الدرجة متمثلا فى التقدير النوعى للأداء وهذا الحكم يفيد فى إتخاذ قرار معين بشأن الفرد الذى حصل على الدرجة أو اقتراح إجراء مناسب له متمثل فى تقدير الأداء النوعى التقدير الوعى للأداء.

جـ- عملية التقييم:

عملية منهجية تتطلب جمع بيانات موضوعية ومعلومات صادقة. من مصادر متعددة بإستخدام أدوات القياس المتنوعة فى ضوء أهداف محدد بغرض التوصل إلى تقديرات كمية وأدلة كيفية يسند إليها فى إصدار أحكام وإتخاذ قرارات مناسبة تتعلق بالأفراد (د/صلاح محمود علام).

سادسا: دور نظرية الاختيار فى الأبحاث والتقييم:

تبدو أهم أدوار نظرية الاختبار من خلال نطاق كبير من الأبحاث وطرق التقييم المنهجى التى تساهم فى تأمل البحث التعليمى والعلوم الاجتماعية كعملية تحقق افضل مراحل البحث المحتمل تصميمها وهذه الأدوار تتمثل فى:

1- صياغة أسئلة البحث أو فروضة.

2- تعيين تعريف جاهز لتقييم الفروض عن طريق تحديد كيفية التحكم والقياس أثناء الدراسة.

3- تطوير اختبار او وسيلة إجرءات نحتاج إليها لتقيس الملاحظة فى كل التقييمات.

4- إيجاد الإختيارات الدقيقة والحساسة فى الوسائل والإجرءات المستخدمة لابد أن تكون واضحة.

5- جمع البيانات التجريبية التى تكون جاهزة للعمل فى التصميمات التجريبية حيث تتيح الفرصة لأسئلة مبتكرة يجاب عنها.

6- إختزال البيانات الرياضية عند تخصيص الإختيارات الإحصائية لتحديد أرجحية نتائج الملاحظة التى تحقق الفرصة المرجوة.

سابعا :التنظيم فى نظرية الاختبار:

التنظيم فى نظرية الاختيار يؤدى لمعلومات حاسمة فى أى بحث أو تقييم. فهو يؤدى لتطور الإختيارات خاصتا النفسية التى تقيس التغير فى الإهتمامات.

التعديل الإستدلالى الأول يقبل فى القياس والاختيارات المستخدمة ، وتطور المجموعات الآخرى من الإختبارات.

سابعا :الخلاصة:

أولا: فى هذا الجزء وحده أساسية مثل متمثلة فى:

أ- إختبار القياس وكيف يمكن تعريفه.

ب- الملاحظات المتشابهة للسلوك ودورها فى تطوير نظرية الاختيار كان يخطط المسافة بين حل المشاكل التى تواجه القياس وتطور بناء القياس.

ثانيا: كما تضم هذه الوحده المشاكل التى وجدت بوجه عام أمام تطور بناء القياس النفسى:

1- لا يوجد تعريف موحد للبناء النفسى تكون عامة ومقبولة.

2- القياس النفسى مبنى على أساس عينة من السلوك.

3- إختيار عينات من السلوك قد تؤدى لأخطاء فى القياس فى أغلب الأحيان.

4- الوحدات فى القياس ليست معرفة بشكل جيد.

5- القياس لابد أن يشرح علاقة التغير الآخر.

ثالثا: خطة نظرية الإختيار:

نظرية الإختيار فرع يرتكز على دراسة كيف يحتمل أن تؤثر هذه المشاكل على القياس النفسى والتوصل للطريقة المثلى للتقليل أو الحد من أثر هذه المشاكل.

رابعا: التطور التاريخى لنظرية الإختبار:

يعرض من خلال تطور الاختيارات النفسية نظرية الاختيار الحديثة تأثرت بالمساهمة فى (المانيا . بريطانيا . فرنسا . امريكا) خاصةً فى علم الاجتماع.

خامسا: نظرية الاختبار تلعب دور هام فى:

العمليات الكلية فى البحث المنهجى عن عرض عام لطرق الاهتمام بالتغير فى القياس وطرق الاختبار الحساسة ودقة فى القياس منظورة الاجراءات.

الفصل السابع

الصدق... المفاهيم الأساسية

إن صدق الاختبار متعلق بما يقيسه الاختبار، وإلى أى مدى يقيس الاختبار ما وضع لقياسة. وذلك يبين لنا أنه يمكن من خلال درجات الاختبار الاستدلال على الصدق. وفى هذا السياق، يجب أن تكون حذرين بخصوص قبولنا لأسم الاختبار كؤشر لما يقيسه الاختبار. حيث أنه عادتا ما يكون هذا الأسم مختصر وواضح مما يسهل مطابقته بالهدف من الاختبار. إلا إننا من ناحية آخرى نجد أن معظم أسماء الاختبارات بعيدة أكثر مما ينبغى عن الوضوح وتكاد تكون مبهمة. بالتالى فهى تمدنا بمعنى غير واضح ومساحة محجوبة من السلوك. ولذلك نجد أن هناك عدد من المحاولات التى تسعى لجعل اسم الاختبار أكثر دقة ووضوحا. من ناحية آخرى نجد أن السمة المقاسة التى يقدمها الاختبار تتضح من خلال الفحص الموضوعى لمصادر المعلومات والعمليات التجريبية المستخدمة فى تقييم صدق الاختبار.

بالإضافة إلى ذلك، نجد أن صدق الاختبار لا يمكن وصفة عن طريق التعبيرات العامة، بمعنى أننا لا نستطيع أن نقول أن هذا الاختبار ذو صدق "عالى أو منخفض" بشكل نظرى. وذلك لأنه لابد من تقييم هذا الصدق وبدقة. وفى الأساس، نجد أن جميع الإجراءات المستخدمة فى تحديد صدق الاختبار متعلقة بالعلاقات القائمة بين الأداء على الاختبار والحقائق المستقلة الأخرى التى لابد من ملاحظتها والتى تدور حول بحث ومعرفة مميزات السلوك، ونلاحظ أن الطرق الدقيقة المستخدمة فى التحقق من هذه العلاقات المتعددة والتى تم وصفها بمسميات مختلفة، هى فى الأساس ترتكز على مختلف مظاهر الصدق

وتنـال إهتمـام خـاص مـن مختلـف مـستخدمى الاختبـارات ومـا هـى إلا محـاولات مستخدمة لتطوير الاختبارات. ووفقا لذلك قد تم تغيير مفهوم الصدق.

تطور مفهوم صدق الاختبار:

منـذ وقـت مبكـر كـان استـخدام الاختبـارات يقتـصر علـى تقـديـر كيفيـة تعلم الأفراد مجالات ذات محتوى محدد. أمـا اليـوم ، فنجـد أن هـذا الاسـتخدام قـد تم الإعلان عنه وظهورة وذلك من خلال طريقة الامتحانات فى المدرسة وكذلك الاختبـارات المتعلقـة بإستخراج رخصة القيـادة أو التأهيـل لمهنـة معينـة. وعمومـا ، نجد أن هذا النوع من الاختبـارات يمكـن تصنيفـه علـى أسـاس أنـه اختبـار تحصيلـى مقييم ، عـن طريـق مقارنـة محتواه مـع محتـوى آخـر مـصمم خصيصا لتقييم ، وذلـك كالمدخل الوصفى الذى مازال يستخدم فى تحديد صدق الاختبارات.

وتنتقل بعد ذلـك إلى مرحلة آخرى ، وهـى التأكيـد علـى التنبؤ. فكيف يختلـف الأفـراد فـى اسـتجابتهم فـى مواقـف معينـة فـى الوقـت الحـالى أو فـى المستقبل؟ وكيـف يـؤثر الشخص فـى المواقـف المختلفـة؟ والأداء فـى هـذا الموقـف يجعلنا نتنبأ بالسلوك. وهذا كله أدى إلى وجود مـا يسمى بـ "المحـك". وبنـاء علـى ذلـك ، يمكننـا أن نقـرر صدق الاختبـار علـى أسـاس الإرتبـاط الـدقيق بـين درجـات الاختبـار وبـين مقيـاس مستقل "المحـك" وهـذا الأجـراء ملائـم بـصورة خاصـة فـى استخدامنا إختبـارات لإنتقـاء الأفـراد فـى بعـض الوظـائف عـن طريـق الـبرامج التربوية أو لبرامج المعالجة الخاصة. مـثلا ، إختبـار الاستعداد الميكانيكى ، المحـك هنـا هـو تنفيـذ العمـل بكفـاءة كالميكـانيكى. وكـذلك إختبـار الإسـتعداد المدرسـى لابد أن يكون لمجموعـة فـى صف دراسـى معـين. أمـا الاختبـارات العـصابية ، لابـد أن تكون التقديرات والمعلومات صادقة فيمـا يتعلـق بسلوك الفـرد فـى مختلـف مواقـف الحياة.

أما المرحلة الحالية فهي تعكس اثنين من أهم الاتجاهات وهما:

1- التوجيه النظري القوى.

2- الارتباط المحكم بين النظرية النفسية والتحقق منها عن طريق التجريب واختبار الفروض. وهذ الاتجاهات واضحة فى بناء الاختبار وتقيمة.

ونجد أن نتيجة هذه الاتجاهات هى التعرف على قيمة التكوين فى وصف وفهم سلوك الفرد ، ذلك لأن التكوين الواضح يشتق من أوجه كثيرة منها الملاحظة المباشرة لمتغيرات السلوك. وهذا النوع من صدق الاختبار يمثل النوع الثالث من أنواع الصدق وهو "صدق التكوين" وهو يبين لنا القاعدة الأساسية لمفهوم الصدق بوجه عام كالدقة فيما يقيسه الاختبار. ونجد أن المحتوى والتنبؤ بالإجراءات الصادقة تتضمن مصادر متنوعة من المعلومات والتى تسهم فى توضيح وفهم تكوين الاختبار. وفى نفس الوقت تزودنا بالمعلومات القيمية التى تجعلنا متفوقين فى تقويم الاختبارات من خلال فهمنا لمفهوم الصدق التكويني.

المحتوى . وصف الإجراءات:

طبيعته:

إن وصف إجراءات المحتوى الصادق تستلزم فى الأساس الفحص المنظم لمحتوى الاختبار. ذلك لتحديد ما إذا كان الاختبار يزودنا فعلا بعينة نموذجية من السلوك المقاس. ونجد أن إجراء الصدق عادتنا ما يستخدم فى الاختبارات المصممة لقياس ومعرفة مدى براعة الفرد فى مهارة معينة أو التقدم فى الدراسة. وذلك يوضح لنا أن فحص محتوى الاختبار لابد أن يكفى لتقييم صدق الاختبار ، مثلا كاختبار المضاعفة (الضرب) والهجاء والتى تبدو أنها تتسم بالصدق عن طريق تحديد محتوى هذه الاختبارات ومفرداتها . ومع ذلك نجد أن الحل ليس

بسيط كما يبدو لنا وأحدى الصعوبات التى تواجهنا تتمثل فى أنه لابد أن تكون عينة المفردات ممثلة وملائمة. كذلك السلوك المحدد للقياس لابد أن يكون محلل بطريقة منظمة وذلك كى يبين لنا أهم المظاهر التى تغطيها مفردات الاختبار، على سبيل المثال، الاختبار المزود بكافة المظاهر عن المجتمع الأصلى يجعلنا أكثر استعدادا لأعداد مفردات موضوعية. وهذا يبين لنا أن ميدان البحث لابد أن يوصف مسبقا وأن يظهر بعد إعداد الاختبار، فمثلا عند بناء إختبار تربوى مقبول، لابد أن تغطى الأهداف التعليمية محتوى الموضوع. وبناء عليه يجب أن يكون المحتوى محدد وشامل لجميع الأهداف. كتطبيق القواعد وتفسير البيانات بالإضافة إلى المعرفة الحقيقية. وعلاوة على ذلك، نجد أن الصدق يعتمد على الاستجابات فى الاختبارات الفردية والتى تكون وثيقة الصلة بالسلوك الذى نقوم ببحثة ودراسته وكذلك بمفردات المحتوى الصادقة والمواضحة. ومن ناحية أخرى نجد أنه يمكن أن يفشل فحص الاختبارات وفى تحديد العمليات المستخدمة فى بحث كيفية إعداد الإختبارات.

كذلك فإن الإجراءات الصادقة هامة فى تحديد أى هدف وفى تعميم الملاحظة على عينة الاختبار فمثلا، اختبار الاختبار من متعدد فى الهجاء فهو يقيس القدرة على التعرف الصحيح على الكلمات وكذلك الهجاء الغير دقيق، وليس معنى ذل أن الاختبار زائف ولكن معناه ان الاختبار يقيس القدرة على الهجاء الصحيح عن طريق تكرار الخطأ فى التهجئة وفى الكتابة وفى مظاهر مختلفة فى القدرة على الهجاء.

وبالرغم من ذلك، مازال لدينا صعوبة تنشأ عن طريق عوامل آخرى لا علاقة فها بدرجات الاختبار، على سبيل المثال ن الاختبار المصمم لقيسا البراعة والتقدم فى الرياضيات أو الميكانيكا يمكن أن يكون له تأثير غير مناسب فى القدرة على الفهم والنجاح فى المهام البسيطة والمكررة.

الإجراءات الدقيقة:

إن صدق المحتوى دائما ما ينشأ بداخل الاختبار، وذلك يحدث فى البداية عن طريق اختيارنا للمفردات المناسبة ، فبالنسبة للاختبارات التربوية نجد أن إعداد المفردات شئ هام جدا يتم عن طريق الشمولية والفحص المنظم للطريقة المناسبة المستخدمة فى تدريس المقرر المدرسى والكتب المدرسية. أيضا عن طريق معرفة أراء الخبراء والمتخصصين. ونلاحظ أن المبدأ الأساسى هنا هو تجميع المعلومات ، وأن مواصفات الاختبار تتضح من خلال المفردة المكتوبة وهذه المواصفات لابد أن تصف المحتوى والموضوعات وتغطى الأهداف التدريسية والعمليات الخاصة بالاختبار. كذلك العلاقة بين الموضوعات والعمليات ، وفى النهاية نلاحظ أن المواصفات يجب أن تشير إلى مجموعة المفردات ونوع الاختبار الذى تنمى إليه هذه المفردات. على سبيل المثال ، إن تقدير القدرة على القراءة يشتمل على فهم كلمات القطعة وكذلك الفهم الحرفى لمحتوى القطعة ، وتصحيح هذه الكلمات نستدل عليه من خلال المعلومات المعطاة ، كذلك يمكن أن تكون عينة المفردات مادة نحصل عليها من مصادر مختلفة ، كالمقالات والقصائد ومقالات الصحف. أما اختبارات الرياضيات فلابد أن تغطى مهارات الحساب وحل المسائل وكذلك تتضمن تحديد مستوى التعلم الجديدة والإجزاء الغير مألوفة.

إن بحث المحتوى الصادق فى تمارين الاختبار التحصيلى التربوى ، يجب أن يتضمن وصف الإجراءات المتبعة والتى تضمن أن يكون المحتوى ملائم ونموذجى ، كذلك لابد أن تتفق أراء الخبراء والمتخصصين فى الحكم على العمليات المكونة للاختبار وفى تصنيف المفردات. وبالرغم من تغيير المنهج ومحتوى الطريقة المستخدمة فى تحديد صدق الاختبار إلا أنها تمدنا بالبيانات وذلك عندما

يتشاور الخبراء والمتخصصين ، وهذ البيانات المعلنة تمدنا بمعلومات عن طبيعة طريقة الدراسة وتقييم الكتب المدرسية.

هناك مجموعة من الإجراءات الاخرى المستخدمة فى تحديد صدق محتوى الاختبار التحصيلى التربوى وذلك عن طريق كلا من مجموع الدرجات والأداء على مفردات الاختبار والتى يمكن أن تستخدم فى توضيح العمليات المكونة للاختبار. كذلك نجد أن هذه المفردات يمكن عن طريقها تحديد نسبة إنتقال الأطفال من مستوى أدنى لمستوى أعلى. بالإضافة لذلك نجد أن الإجراءات لابد أن تتضمن تحليل لكافة أنواع الخطأ الموجود فى الاختبار ، وأيضا ملاحظة طرائق العمل المستخدمة فى الاختبار. وأخيرا ، لابد أن نقوم بإختبار الطلاب كلا على حدى وذلك عن طريق تعليمهم طريقة "التفكير بصوت مرتفع" أثناء حل المسألة. حيث أنه يمكننا عن طريق سرعة مشاركة الأفراد أن نحدد عدد الأفراد الذين فشلوا فى إنهاء الإختبار. وكذلك نكشف عن الأسباب الغير متصلة بموضوع الاختبار والتى تؤثر على قدرة الفرد فى تعلم القراءة من خلال الاختبار المناسب والأداء على الاختبار ، ومن الممكن أن تكو درجات الاختبار مرتبطة مع درجات إختبار القدرة على فهم القراءة. وعلى صعيد آخر ، كان الاختبار صمم لقياس القدرة على فهم القراءة فنجد أنه من خلال الأسئلة المعطاه وبدون قراءة الفقرة سوف توضح لنا عدد الأسئلة التى تم الإجابة عليها ببساطة وكذلك تحدد لنا الأسباب الغير متصلة بموضوع ومحتوى الإختبار.

التطبيقات:

إن المحتوى الصادق يعد طريقة مناسبة لتقييم إختبارات التحصيل ، خاصة عندما يتم تدعيمها بالتحقيق والضبط التجريبى ، كما هو موضوح

مسبقا. كذلك فهو يتيح لنا فرصة الإجابة على سؤالين وهما بمثابة قاعدة أساسية لصدق اختبارات التحصيل التربوى والمهنى.

1- هل يغطى الاختبار عينة نموذجية من المعارف والمهارات المحددة؟

2- هل هذا الاختبار ملائم ومستقل عن العوامل الآخرى الغير متصلة بالموضوع المقاس؟

كذلك فنجد أن المحتوى الصدق يكون ملائم عن طريق المراجعة الخاصة للاختبارات ، وهذه الملائمة تعنى وبشكل واضح أن المحتوى الصادق هو شرط أساسى وأولى عند استخدامنا للاختبارات إلا انا نجد انه انواع آخرى من الصدق مناسبة جدا فى عملية التقييم الكامل لفعالية الاختبارات.

كذلك فإن صدق المحتوى ملائم أيضا للاختبارات المهنية والتى صممت لاختيار وتصنيف الأفراد. وهذا النوع من الصدق يكون مؤشر مناسب خاصة عندما يكون الاختبار واقعى وحقيقى فى تحديد عينة الوظيفة أو بمعنى اخر تحديد المهارات والمعارف المتطلبة فى الوظيفة ، ومثل هذه الحالات ن التحليل الشامل للمهمة والذى لابد أن يتعد كى يظهر لنا بوضوح الشبة بين انشطة الوظيفة والاختبار (ذلك موضح خطوة بخطوة فى تقرير عن استعمال الإجراءات الصادقة لتطوير اختبار القراءة للعاملين وهو موضح عن طريق Schoen (Feldt, Schenfdldt, Acker and Perlson (1976) ونجد أن العمل النهائى الذى يقوم به أصحاب الوظيفة والمشرفين هو التحقق من التقارير وقراءة متطلبات المستوى المدخلى للوظائف وتوضيحها فى الشركات الصناعية الكبرى.

كذلك فلابد أن تكون مفردات الاختبار ذات تكوين يسمح لها باختبار هذه المتطلبات. وهذه الطريقة تستخدم بتوسع فى تطوير الاختبارات المصممة للعاملين فى الحكومة وفى الاتحاد الفيدرالى ومستويات الولاية بأكملها.

على صعيد اخر، نلاحظ أن صدق المحتوى عادتا ما يكون غير مناسب وخادع فى اختبارات الشخصية والاستعداد. وقد دلت كثير من الأبحاث أن المحتوى النموذجى لابد أن يكون واضح من خلال مراحل التكوين الأولى للاختبار. وبذلك نجد ان كى نقوم بتقييم الصدقفى اختبارات الاستعداد والشخصية لابد من التحقق من الإجراءات فلابد أن تكون واضحة ومتسقة. كذلك لابد أن تقودنا هذه الاختبارات لصورة واضحة عن السلوك المراد قياسة. وبناء على ذلك، نجد أن محتوى اختبارات الشخصية والاستعداد يمكن أن يكشف لنا فروض الاختبار والتى توضح لنا نوع السمة المقاسة، وهذه الفروض تحتاج لإثباتها تجريبيا حتى نتمكن من تقييم صدق الاختبار.

وبخلاف الاختبارات التحصيلية، نجد أن اختبارات الاستعداد والشخصية ليست أساسية فى تحدي سير التعليمات والضبط المتسق للتجارب السابقة والتى تظهر فى محتوى الاختبار. ومن هنا، نجد انه فى الاختبارات الثانية (الاستعداد والشخصية) يمكن تحديد الأفراد الملائمين فى كثير من طرائق العمل أو فى العمليات التقسية المستخدمة فى الاستجابة على نفس مفردات الاختبار. وبالتالى، فإن نفس الاختبار لابد أن يقيس مختلف الوظائف لمختلف الأفراد. كذلك يمكن أن تكون هذه الاختبارات فعالة فى تحديد الوظائف النفسية المقاسة عن طريق فحص محتوى الاختبار، على سبيل المثال، أن خريج الكلية يجب أن يحل المسائل سواء فى صورة لفظية أو مصطلحات حسابية، بينما الميكانيكى يمكنة تقديم نفس الحل فى مصطلحات عن التخيل المكانى. أو الاختبار الذى يقيس الاستنتاج الحسابى بين الطلاب المبتدئين فى المدرسة العليا فهو يقيس فقط الفروق الفردية فى سرعة الفهم عندما يقدم لطلاب الجامعة.

الصدق الظاهرى:

إن صدق المحتوى لا ينبغى أن يكون مشوش بما تسمية الصدق الظاهرى ، ذلك لأن الصدق الظاهرى ليس صدقا بالمعنى التقنى ، لأنه لا يشير إلى ما يقيسه الاختبار فعلا ، ولكنة يشير إلى ميا يبدو أن يقيسه الاختبار بشكل سطحى. ونلاحظ أن الصدق الظاهرى يتعلق بأحد الأثنين وهما الاختبار الذى "يبدو صادقا" للمتحنين المستخدمين هذا الاختبار ، والأشخاص التنفيذين والمراقبين النفيين الذين يجسموا هذا الاستخدام. وأساسى ، نجد أن سؤالنا عن الصدق الظاهرى لابد أن يتعلق بمدى الألفة والعلاقات العامة. ذلك وبالرغم من الاستعمال الشائع لمفهوم "الصدق" إلا أنه يمكن أن يؤدى بنا إلى وجود إرتباك وتشويش. فعلى سبيل المثال ، عندما صممت الاختبارات فى الأصل للأطفال ، وتطورت بعد ذلك داخل محيط حجرة الدراسة واستخدمت بشكل موسع معالراشدين. فكثير ما واجه ذلك مقاومة وإنتقاء بسبب فقدان الصدق الظاهرى. ومن غير ريب ، أنه إذا أظهر أن محتوى الاختبار غير متصل بالموضوع وغير ملائم وبسيط ، فستكون النتيجة أن هذا الاختبار ردئ ومهيل لصدق الاختبار. خاصة فى اختبار الراشدين ، فنجد أنه غير كافى لجعل الاختبار ذوى صدق موضوعى. كذلك فإننا أيضا نحتاج الصدق الظاهرى لتحديد فعالية العمل فى مواقف عملية مختلفة. ومن ناحية آخرى نجد أن الصدق الظاهرى يؤثر على مقبولية الاختبارات فى السلطة والقرارات التشريعية ، بالإضافة لذلك يمكننا أن نقييم الصدق الظاهرى عن طريق فكرة عامة عن الاختبار.

وقد اوضح (Baruch Nevo) ومساعدية ان التحقق من الصدق الظاهرى يتضمن برنامج بحثى إبتكارى عند إجراء الاختبار ، وهو بمثابة فحص يتم عن طريق الاختبار المطبق. ونجد ان اول شئ استدعى إنتباههم هو قلة الأبحاث فى الصدق الظاهرى. كذلك فهم اقترحوا تقديرات كمية للصدق الظاهرى وذلك يتم بإعطاء الأفراد درجة ملائمة على الاختبار المستخدم ونفس

الإجراءات يمكن أن تستخدم فى تصنيف وتقدير مفردات الاختبار الفردى أو بطارية الاختبارات.

وغالبا ، ما يمدنا الصدق الظاهرى بصياغة لمفردات الاختبار فحسب ، وذلك فى هيئة تغييرات تظهر مناسبة معقولة ظاهريا على سبيل المثال ، إذا كان لدينا اختبار للتفكير الحسابى البسيط ، نجد أن المفردات لابد أن تصاغ فى مصطلحات وتعبيرات خاصة بالعمليات الحسابية ومثال لذلك كم عدد البرتقال الذى يمكن شرائه بـ 86 سنت؟ أو أى مثال آخر من مسائل الكتاب المدرسى التقليدى. وبالمثل ، نجد أن الاختبار الحسابى للأشخاص الذين يعيشون فى المناطق البحرية يمكنة أن يظهر مصطلحات فنية خاصة بهم ، وذلك بدون وجود ضرورة لتعديل وتغيير الوظائف المقاسة. ونؤكد أن الصدق الظاهرى لابد أن تغطية نفس الاهتمام الخاص بالصدق الموضوع.

المحك - التنبؤ بالإجراءات:

الصدق التنبؤى والتلازمى:

إن محك التنبؤ بصدق الإجراءات يعبر بإختصار عن فاعلية الاختبار فى التنبؤ بمدى ملائمة الأفراد فى أنشطة محددة. لذلك نجد أن درجة المحك فى المقابل لابد أن تعبر عن درجات الاختبار الصادقة التى يتم الحصول عليها بشكل تقريبى ، وذلك يتم أما فى نفس وقت تقدير الاختبار أو بعد فترة فاصلة من التطبيق. وفى عام (1985) قد تم اختيار بعض القواعد الأساسية المميزة بين الصدق التلازمى والتنبؤى وذلك فى ضوء العلاقات القائمة بين المحك والاختبار. ونجد أن مصطلح "التنبؤ" يمكن استخدامة بمعنى أعمق وذلك عندما يشير إلى التنبؤ عن طريق محك وظيفى. أو بمعنى أكثر تعقيدا التنبؤ بواسطة وقت فاصل. وذلك كلة متضمن فى مصطلح "الصدق التنبؤى" ونلاحظ ان

المعلومات التى يتم الحصول عليها بواسطة الصدق التنبؤى وثيقة الصلة بالاختبارات التى تستخدم فى اختيار وتصنيف الأفراد ، كذلك يستخدم عندما نحتاج بعض الأفراد للعمل فى وظيفة ما. واختيار الطلاب للقبول بالجامعة أو بالمدارس المهنية والحرفية كذلك فى برامج التدريب المهنى للأفراد الذى يتم تعينهم فى القوات المسلحة والتى توضح لنا أمثلة لعدد من القرارات التى تحتاج إلى معرفة الصدق التنبؤى للاختبارات. وهناك أمثلة آخرى تتضمن استخدام الاختبارات للكشف عن الاضطراب والهياج الإنفعالى الناتج عن ضغط البيئة المحيطة ، كذلك استخدام الاختبارات من قبل الطب النفسى فى مساعدة المرضى وأمدادهم بالعلاج المناسب.

وفى عدد من الحالات ، نجد أن الصدق التلازمى يستعمل فحسب كبديل عن الصدق التنبؤى والذى كثيرا ما يكون غير عملى فى إبراز وتقييم صدق الإجراءات مع مرور الوقت المنقضى أو للحصول على عينة ملائكة من أهداف الاختبار. ولكنة يمثل محك لبيانات متاحة فى ذلك الحين. وهنا نجد أن درجات أختبار الطلاب فى الجامعة يمكن أن تقارن عن طريق درجة محددة تمثل محك.

وعلى صعيد آخر ، نجد أن أنواع الصدق سواء كانت (تلازمى . تنبؤى) تعتبر أكثر الأنواع ملائمة ويمكن عن طريقة إثبات صحة الاختبار. وذلك يظهر لنا فى الاختبارات النفسية. ونلاحظ أن الاختلاف المنطقى بين الصدق التنبؤى والصدق التلازمى لا يتمثلا فى الوقت فقط ولكن أيضا فى الهدف من أجراء الاختبار. فنجد أن الصدق التلازمى مناسب فى الاختبارات المستخدمة من أجل تشخيص الوضع الحالى للسلوك أو على الأصح التنبؤ بالمستقبل القريب. ويمكن ان نوضح هذه الاختلافات عن طريق الاسئلة التالية:

1- هل سميث مؤهل كقائد بارع؟

2- هل سميث لدية المتطلبات الأساسية التى تجعله قائد بارع؟ إن أول سؤال يعبر عن الصدق التلازمى أما السؤال الثانى فهو يبين لنا الصدق التنبؤى.

تلوث المحك:

عن إتخاذ الحذر الكامل عند اكتشاف صدق الاختبار، يمكننا من التأكيد على أن درجات الإختبار ليس لها تأثير على وضع الأفراد فى المحك. على سبيل المثال، إذا كانت تقديرات الأفراد الخاصة بإختبار الإستعداد ضعيفة جدا. فإن معرفتنا بذلك يمكن أن يؤثر على المستوى الذى سوف نحدده للطلاب أو التقدير الخاص بالعمال. أما التقديرات المرتفعة للأفراد فيمكن أن تؤدى لحدوث شك وارتياب عندما نقوم بمقارنتهم بمستوى أكاديمى معين أو التقديرات الوظيفية. وهذه التأثيرات يمكن أن تؤدى إلى إرتفاع الارتباط بين درجات الاختبار والمحك. والذى بدورة يكون زائف وغير حقيقى.

مصدر الخطأ الذى يمكن أن يظهر فى صدق الاختبار يمكن أن نعرفة بـ "تلوث المحك" ونجد أن تقديرات المحك يمكن أن يحدث لها تلوث أو فساد عن طريق معرفة تقديرات درجات الاختبار. وكى تتخطى فعالية وتأثير اى مصدر للخطأ، يجب ان نتأكد وبشكل اساسى أن الفرد الذى يشترك فى مهمة "تقديرات المحك" لا يعرف الدرجات الخاصة بالاختبار. لذلك السبب نجد أن درجات الاختبار التى يتم استخدمها فى "إختبار الاختبار" لابد ان يقيم على نحو خاص وبشكل سرى. وذلك أحيانا ما يكون صعب بالنسبة لأقناع كل من المدرسين وأصحاب العمل وضباط أو موظفين القوات المسلحة بذلك. ومن ناحية آخرى نجد أشخاص أكثر إلتزام وحذر عند أصدار القرارات. حيث انهم لا يتأثر بدرجات الإختبار ولكن يضعوها جانبا حتى يمكنهم دراسة بيانات المحك والتحقق من الصدق.

مقاييس المحك (أنواع المحكات):

أن الاختبار يمكنة أن يكون صادق فى ضوء عدد من المحكات المستخدمة لهذا الفرض. وأى طريقة يتم استخدامها لتقدير السلوك فى اى موقف لابد أن يكون لدينا محك للتحقق من الأهداف الخاصة. ونجد أن المحكات التى يتم استخدامها فى الكشف عن الصدق تكون موضحة فى تعليمات الاختبار، وهى تنقسم إلى فئات وأنواع عامة، ومن ضمن المحكات المستخدمة كثيرا لتحديد صدق اختيارات الذكاء (التحصيل الدراسى) حيث أننا نجد أن كل الاختبارات غالبا ما يتم وصفها كمقاييس تشتمل على درجة الاختبار التحصيلى، التشجيع، درجات الشرف، الجوائز، تقديرات المدرسين وأساتذة الجامعة عن (الذكاء). وكل التقديرات المحددة داخل المحيط الدراسى مناسبة لتحديد أداء الفرد المدرسى والذى يمكن أن نيستخدمه كمحك للتحصيل الدراسى.

ونجد أن المؤشرات المختلفة للتحصيل الدراسى توضح لنا بيانات المحك فى كل المستويات التعليمية، من خلال المستويات الأساسية للأفراد. وبالرغم من استعمال هذا المحك بالدرجة الأولى فى صدق اختبارات الذكاء العام، إلا أننا ايضا يمكن أن نستخدمه كمحك لاختبارات الشخصية، والاستعدادات المتعددة. وعن طريقة تحديد صدق أى نوع من هذه الاختبارات يمكننا أن نختار الطلاب للإلتحاق بالجامعة.

ونلاحظ ان محك التحصيل الدراسى قد أختلف كثيرا عندما قمنا بإستعماله بعيدا عن المدرسة وذلك من خلال معرفة مقدار التعليم الذى حققه الأفراد قبل ترك المدرسة. ومنه نستطيع أن نتوقع أن الأفراد الأكثر ذكاء قد قضوا مدة أطول فى التعليم بينما الأفراد الأقل ذكاء قد تركوا المدرسة في وقت مبكر.

فـى تطـوير اختبـارات الاسـتعداد الخـاص ، يوجـد نـوع آخـر مـن المحكـات يعتمد على (الأداء فى التدريب التخصصى) على سبيل المثال. اختبارات الاستعداد الميكـانيكى يمكـن أن تكـون صـادقة فـى ضـوء التحـصيل النهـائى فـى المدرسـة الصناعية. وتتنوع مقررات المدارس المهنية مثل الكتابة على الألة الكاتبة ومسك الـدفاتر والتى تعتبر محكات لكثير مـن اختبـارات الأستعداد فى هـذه المجـالات ، وبالمثـل ، الأداء فـى مـدارس الموسـيقى يـستخدم كمحـك لـصدق اختبـارات الاستعداد الفنى أو الموسيقى. كذلك فإن اختبارات الاستعدادات المهنية يمكن أن تكون صادقة من خلال التحصيل فى مدارس الحقوق ، الصيدلة ، طب الأسنان ، والهندسة وكذلك اى مجالات آخرى. ونلاحظ أن صدق اختبـارات إنتقـاء قائـدى الطـائرى العسكرية يتم فى ضوء الأداء وهو التدريب على الطيران ونجد ان هـذا المحك يمكن استخدامه لتحديد صدق الاختبـارات المهنيـة العسكرية وفى بعض الدراسات الصناعية الصادقة.

ومن بين المؤشرات الخاصة بالأداء فى التدريب التخصصى ، نجد أن الأداء المـستخدمة مـن خـلال أهـداف المحـك يمكـن أن يـشير إلى اختبـارن تحـصيلية مستخدمة فى إكتمال التـدريب وكـذلك تقـديرات المعلمـين والـدرجات المحـددة بـشكل صـورى. امـا بخـصوص بطاريـات الاسـتعداد المتعـددة ، فإنهـا غالبـا مـا يـتم فحصها فى ضوء الدرجات فى المدرسة الثانوية أو مقررات التعليم الجامعى. ونقوم بتحديـد صـدقهم كمنبـئين مخـتلفين. علـى سـبيل المثـال. الـدرجات فـى اختبـار الفهم اللفظي يمكن أن يتم حسابها عن طريق الدرجات فى اختبار الفهـم اللفظي يمكـن أن يـتم حـسابها عـن طريـق الـدرجات فـى مقـرر اللغـة الإنجليزيـة ودرجـات التخيل المكانى ودرجات علم الهندسة وعلم جرا.

فـى سياق اسـتخدامنا لـدرجات وتقـديرات التـدريب بشكل عـام كمحـك للقياس ، فإننـا يمكننـا ان نميـز بـين المحكات الوسـطى والنهائيـة فـى تطـوير اختبـار

لانتقاء قائد سلاح الطيران أو اختبار للاستعداد الطبى ، على سبيل المثال ، إن المحكات القصوى (النهائية) تقوم بتحديد الأداء النهائي أو التحصيل النهائي كالكفاءة الجراحية للطبيب ، ونلاحظ أن ذلك يتطلب وقت طويل لدراسة بيانات المحك ، مما يجعله مبهم ويصعب قياسه ، ونجد أنه يمكن أن يتعرض فى النهاية لكثير من العوامل التى يصعب التحكم فيها والتى تؤثر على استخدامه. على سبيل المثال ، نجد أنه من الصعب تقدير درجة النجاح فى مهنة الطب بمختلف تخصصاتها وأقسامها . ولهذه الأسباب كثيرا ما تستخدم المحكات الوسطى لتسجيل الأداء فى بعض مراحل التدريب وكثيرا ما تستخدم كمقاييس مرجعية.

لعدد من الغايات ، نلاحظ أن معظم تصنيفات المقاييس المحكية تعتمد على تسجيل مستوى الأداء الفعلى فى العمل. ومحك (تسجيل الأداء الفعلى فى العمل) يستخدم فى نطاق صدق اختبارات الذكاء العام ، بالإضافة إلى اختبارات الشخصية واختبارات الاستعداد الخاص. وهذا المحك بوجه عام يستخدم فى تحديد صدق الاختبارات التى تم تصميمها لبعض الوظائف الخاصة ، وهذه الوظائف عندما يتم طرحها فى أسئلة يمكن أن تتنوع فى كلا من المستوى والنوع فتتضمن العمل فى التجارة ، الصناعة ، الحرف ، الخدمة العسكرية. ومعظم مقاييس الأداء الوظيفى بالرغم من ذلك لا تحدد كمحكات قصوى (نهائية) ولكن تمدنا بمحكات وسطى من خلال اختبار عدد من الأهداف. نظرا لذلك قد تم تفصيل ما يسمى بـ (سجلات التدريب) على صعيد آخر ، نجد أن مقياس الأداء الوظيفى لا يتيح الفرص المنظمة لتماثل الحالات خلال التدريب. ونلاحظ أن تعليمات الاختبار يمكن أن تصف لنا بيانات الصدق فى مقابل محكات وظيفة تصف المقاييس المحكية المستخدمة بجانب المهام الوظيفية التى يتم إنجازها عن طريق العمال.

الصدق بواسطة طريقة (المجموعات المتناقضة) إن هذا المحك محك مركب يعكس لنا الاختبار الغير متحكم فيه والذى يظهر ويؤثر على حياتنا. على سبيل المثال، صدق اختبار الاستعداد الموسيقى واختبار الاستعداد الميكانيكى يمكن التحقق منه عن طريق مقارنة درجات الطلاب التى تم الحصول عليها فى مدرسة الموسيقى او فى مدرسة الهندسة مع درجات طلاب أحدى المدارس الثانوية أو طلاب الجامعة. وبالطبع نجد أن نحك المجموعات المتضادة يمكن استخدامه، كالتقديرات والأداء الوظيفى. والمجموعات المتضادة تتضمن فئة من الأفراد، يمثلوا مجموعات متميزة وواضحة تتدرج لتصبح مختلفة من خلال اعمال متعددة فى حياتنا اليومية.

إن طريقة المجموعات المتناقضة تستخدم عمومات فى تحديد صدق اختبارات الشخصية. وفى حساب صدق اختبار السمات الاجتماعية، كأداء الشخص الذى يقوم بالبيع أو المديرين على الاختبار، يمكن مقارنته من خلال رجال الدين والمهندسين. ونجد أن الافتراض الأساسي فى هذا الأجراء هو أن الأفراد الذين يباشروا أعمالهم يمثلوا مجموعة من الأفراد فى مجال معين كعمل الكاهن أو المهندس وبالمثل طلاب الجامعة المنشغلون بالأنشطة اللاصفية يمكن مقارنتهم مع طلاب ليسوا مشتركين فى أى نشاط داخل الجامعة. وقد تم استخدام هذا المحك فى تطوير وحساب صدق اختبارات الميول، مثل (Svib) وكثير ما يتم استخدام هذا المحك فى حساب صدق مقاييس الاستعداد.

كذلك فإننا يمكن أن نستخدم طريقة المجموعات المتناقضة فى تحديد الصدق التجريبى لبعض الاختبارات. ولذلك، نجد أن المجموعات المختلفة فى مقدار التعليم يتم مقارنتها فى اختيار ملائم. وإذا كانت تقديرات التفوق قد تم إعطائها، فيمكننا أن نعمل تحليل (2×2) بمعنى أننا نوضح نسبة درجات النجاح

والفشل فى مجموعة التعليم القبلى ونقارنها بنسبة درجات النجاح والفشل فى مجموعة التعليم البعدى.

وفى تطوير اختبارات الشخصية ، يمكننا أن نستخدم التشخيص النفسى كأساس لاختبار مفردات الاختبار وكدليل على صدق الاختبار. كذلك فإن التشخيص النفسى يمكن تقديمه كمحك مرض يعتمد على الملاحظة المستمرة وعلى معرفة تاريخ الحالة والمقابلة السريعة إلا أننا بالرغم من ذلك لا يمكننا أن نعتمد عليه كمحك للقياس ولكننا يمكن أن نستخدمه كمبنئ لتحديد الصدق.

ويمكننا أن نشير إلى نوع آخر من أنواع المحكات يمكن أن نستخدمه وهو التقديرات التى نحصل عليها من خلال مدرسى المدرسى ووجهات النظر الخاصة بالمعلمين ومشرفى الوظائف وكذلك فإننا يمكننا جميع تقديرات عن الضباط فى الجيش وعدد الطلاب فى المدرسة وأيضا تقديرات عن العاملين ، والتى تمثل أسلوب يساعدنا فى الحصول على معلومات بخصوص عدد من المحكات كالتحصيل الدراسى والأداء فى التدريب التخصصى أو النجاح الوظيفى. والآن يمكننا أن نستخدم التقديرات كوسيط لقياس المحك وفى هذه الحالات ، نجد أن التقديرات توضح المحك ، بالرغم من أنها ليست مقتصر على تقييم التحصيل الدراسى ولكن تستخدم فى الحكم على الأفراد عن طريق الملاحظة المتصلة بمختلف السمات التى يقسها الاختبار النفسى ونجد أن صدق العينات يمكن أن يشترك فى تقييم كل الصفات المميزة للفرد ، كالإبداع الميكانيكى والأصالة ، والقيادة ، وكذلك الصدق.

ونلاحظ أن التقديرات يمكن أن تستخدم فى تحديد صدق كل الاختبارات ، كذلك فهى ناقصة فى أنها تمدنا بمحكات عن اختبارات الشخصية. فى حين أن المحكات الموضوعية تعمل على تحديد السمات الاجتماعية بوضوح ،

والتقديرات تعتمد فى الغالب على صعوبة كبيرة لإيجادها فى هذا المجال. وذلك فى الحقيقة خاص بالسمات الاجتماعية الواضحة ، والتى يعتمد تقديرها على الاتصال المباشر بالفرد. وبالرغم من أن التقديرات يمكن أن تخضع لكثير من اخطاء الحكم. إلا اننا يمكن أن نحصل عليها فى حالات متحكم فيها عن طريق استخدام أساليب لتحسين الدقة فى التقديرات وتقليل الأنواع المختلفة من الأخطاء (موضح فى الفصل الـ 16).

فى النهاية ، نلاحظ أن الارتباطات بين الاختبار الجديد والاختبارات المتاحة لدينا تمكن أن تمثل دليل على صدق هذا الاختبار. وذلك عندما يكون الاختبار مختصر ودقيق. فإذا نظرنا إلى اختبار الورقة والقلم ، فإننا نجد أنه يمكن تقدير صدقة فى ضوء أشياء كثيرة متحكم فيها كالوقت المستهلك عند أداء اختبار. كذلك فنلاحظ أنه يمكن أن يكون اختبار الجماعة صادق فى ضوء اختبار فردى ، على سبيل المثال ، أختبار (ستاتفورد . بنية) قد تم استخدامه كثير كمحك لحساب صدق عدد كبير من الاختبارات الجماعية.

نجد انه فى عام 1980 - 1990 قد حدث تطور هام فى بناء الاختبار ، والذى بدوره ركز إنتباهنا على تحليل المحك ، وذلك يمثل احد مظاهر تطوير الاختبار الذى تم أهماله بشكل عام فى الأبحاث التقليدية للاختبارات. ومع مرور الوقت ، نجد أن هناك أصوات نافذة قد طالبت بالبحث النظامى عن المحك وتحليله ولكن هذا الإنجاز كان ضئيل (L.R. James, 1973, Tenopyr, 1986) اما فى وقتنا الحالى ، فقد بدأ فى تعريقنا بشكل كبير بأن صدق الاختبار يمكن أن يكون فعال فى البحث عن تماثل التكوينات فى اداء الوظيفة او المهمة المعطاه وفى اختبار وتطوير الاختبارات التى يتم تقدير درجاتها على أساس حاجاتنا للتكوين (J. P. Campbell, 1990a, J.P. Campbsll, Mc Henry & Wise, 1990, L.V. Jones & Appelbaum, 1984

(Messick, 1995) وفى مثال لتحديد تطبيق البحث المتعلق بالمحك ، أول خطوة فى ذلك هو تطوير بطارية الاختبارات التى أمدتنا بها الولايات المتحدة حول الاختبار فى الجيش وخطط التصنيف وتعرف بـ (خطة A) (J. P. Campbell, 1990b) وهو متصل باستخدام الاختبارات فى الصناعة وفى التنظيم.

تعميم الصدق:

إن الصدق المحك التنبؤى ، غالبا ما يستخدم فى دراسات الصدق المحلية ، وتأثيرتها على الاختبار من خلال برامج خاصة لتقيمها وتحديدها ، وهذا هو المدخل المستعمل ، على سبيل المثال ، عندما ترغب شكرة معينة فى تقييم اختبار لإنتقاء الأفراد لواحدة من الوظائف ، أو عندما ترغب الجامعة فى تحديد الأفضل من خلال اختبار للإستعداد الإكاديمى ، يمكننا أن نتنبأ بالأداء المتقدم للطلاب ، أى أن صدق المحك يتم وصفة كصدق عملى للإختبار من خلال أهداف محددة.

ونلاحظ أن اختبارات الاستعداد المقننة تكون مرتبطة أولا مع الأداء على مهام متشابهة فى دراسات صادقة للصناعة ، بالرغم من أن معاملات الصدق كانت متجهة إلى تنوع كبير (Ghiselli, 1959, 1966) وبالمثل التغيير فى معاملات الصدق يكون ملاحظ عندما تكون المحكات مصنفة فى مقررات دراسة مختلفة (G. K. Bennett, Seashore, Wesman, 1984) ونجد أن كل النتائج تمثل بنا إلى التشاؤمية فيما يتعلق بإمكانية تعميم الصدق فى مختلف الأماكن. وفى منتصف عام 1970 قد تم (تحديد الموافق) والحاجات النفسية ، فيمايتعلق بالقصور فى استخدامات الاختبارات المقننة فى إنتقاء الأفراد ، وفى التحليل الإحصائى للمشاكل ، قد قام كل من Hunter and Schmidt ومساعديهم بتوضيح كمية التباين بين معاملات الصدق التى نحصل عليها

والتى يمكن أن تكون نتائج إحصائية صناعية ناتجة عن عينة صغيرة الحجم ، محك غير ثابت ، وتقيد وحصر لمدى عينات العاملين.

ونجد أن عينات العمال المتاحة من خلال صدق الاختبار تكون صغيرة جدا كى تتح تقدير مستقر/ثابت للعلاقة بين المنبئ والمحك. ولنفس السبب ، نجد أن معاملات الصدق التى نحصل عليها تكون منخفضة جدا لتمدنا بدلالات إحصائية عن العينة المستعملة وربما تفشل فى أمدادنا بدليل على أن الاختبار صادق ، والذى يتم تقديره عن طريق تصيف عينات الصدق المستخدمة فى الدراسات الصناعية والتى لا تشمل على أكثر من 40 أو 50 حالة (Schmidt, Hunter Surry, 1976) عن طريق العينات الصغيرة ، نجد أن تطبيقات الصدق المرتبطة بالمحك التنبؤى ليست تقنية أو عملية.

ونلاحظ أن استعمال الأساليب المطورة الجديدة للبيانات من خلال عينات كثيرة يظهر لنا من خلال مجموعة كبيرة من الخصائص للمميزات المهنية (Hunter, Schmidt) ومساعديهم ، قد قاموا بوصف وتوضيح أن صدق اختبارات الاستعداد (اللفظى ، العددى ، الاستنتاجى) يمكن أن يعمم على نحو واسع فى المهن المتعارف عليها بوضوح. وأن اختلاف معاملات الصدق التى وجدت فى الدراسات الصناعية المبكرة قد اثبتت بأنها ليست افضل من المتوقعة بواسطة الخطأ الصدفة. وذلك صحيح حتى عندما تكون المهام الأداء الناجح لمدى التغيير فى المهات الوظيفية يعتمد على مستوى دلالة مقبولة فى صميم المهارات المعرفية المشتركة. ونجد أن الاختبارات المتضمنة فى هذه الدراسات تغطى نوع المحتوى والمهارات المختبرة خاصة فى اختبارات الذكاء التقليدية واختبارات الاستعداد المدرسى وذلك يبدو وكأنه عنقود من المهارات المعرفية التى تعتبر تنبؤ واضح للأداء فى كلا من الأنشطة المهنية والتحصيلية الأكاديمية المتطلبة فى المجتمعات المتقدمة. بالرغم من ذلك نجد أن كثير من القرارات المتعلقة بنسب

الاختبار يمكن إثباتها إذا أخذنا بعين الاعتبار الدرجات فى اثنين أو ثلاث مجموعات معرفية واضحة. وعن طريق مقايين تقنية للمهمة خاصة المهارات الخاصة للوظائف (Hatrigan, Wigdor, 1989, L. L. Wise, Mchenry, .(Campbell, 1990, Zeidner, Johnson, 1991).

التحليل البعدى:

إن الإجراءات الإحصائية المستخدمة فى التحقق من تقييم الصدق تمدنا فى الأساس بنتائج كاملة من مختلف الدراسات. ويمكننا أن نجمع البيانات من خلال الأبحاث المتعلقه بالماضى والحاضر. من خلال الدراسات المواكبة لمختلف الأماكن أو م نخلال المعلومات المتاحة فى الدراسات المعلنة. بالرغم من أول إشارة لتعميم (التحليل البعدى) قد قدمت فى البحث عام 1970 (Glass, 1976; Schmidt, Hunter, 1977). إلا أن إجراءاته الأساسية قد تم استخدامها خلال عقود كاملة، وخاصة فى العلوم الأخرى (Lipsey, Wilson, 1993; Schmidt1 992). كذلك فنجد أن المسح النموذجى لمجموعة من الدراسات قد اظهر وجود تأثيرات إحصائية دالة. على سبيل المثال، الاختلافات ما بين مسعنى المجموعة الضابطة والتجريبية أو الارتباطات بين درجات الاختبار والمتغيرات الأخرى. عن طريق هذه الإجراءات يمكننا أن نستدل على نتائج إيجابية غالبا ما يتم إهمالها نتيجة لأن العينات المستخدمة فى الدراسة عينات صغيرة جدا مما يؤدى لظهور اختلافات فى الدلالة.

وعن طريق تجميع نتائج ابحاث الدراسات المستقله وإمكانية ارتباطها بالمعالم الأساسية للدراسة. فنجد أن (التحليل البعدى) يكشف لنا عن نتائج حقيقية وهامة. وفائدة أخرى للتحليل البعدى وهو انه يتيح الفرصة لتقدير حجم التأثير فى كلا من الهدف النظرى والعملى. كذلك استنتاج حجم

الارتباطات والاختلافات غالبا ما تكون مفيدة حتى فى الشرح البسيط لأنه ذو مستوى دلالة اكبر من الـ Zero.

وفى العقد (1980- 1990) قد شهد زيادة سرعة فى الأبحاث المتعلقه بـ (التحليل البعدى) فى معظم حقول علم النفس. إن تطبيقاته فى الاختبار المهنى وتصنيف البحوث تلفت انتباهنا (فصل 17) كذلك الاهتمام بالتحليل البعدى فى زيادة مستمرة وذلك نتيجة للتحسينات الإجرائية المطورة. بالرغم من ان هناك جدل قائم حول الأساليب الخاصة والنتائج الرئيسية الغير مختلفة عن تقدير الإجراءات المستعرضه.

التكوين. معرفة الإجراءات:

إن مصطلح (صدق التكوين) قد تم الإعلان عنه وتقديمه بصورة رسمية عن طريق العالم (Lexicon) وذلك عام 1954 فى التوصيات الفنية للاختبارات النفسية وأساليب التشخيص. والذى قدم أول نسخة لمستويات الاختبارات الشائعة ، وأول شرح منفصل لصدق التكوين قد ظهر فى السنة التالية فى مقال لـ Cronbach and Meehl وذلك عام 1955. والمناقشات التى دارت حول صدق التكوين والتى تتابعت وتواصلت قد جاءت لتوضيح الإجراءات المحددة والاستخدام النظامى المنطقى لها. إن مفهوم صدق التكوين يوجه انتباهنا إلى أهمية دور النظريات النفسية فى تكوين الاختبار وفى صياغة الفروض ومعرفة المثبت منها والغير مثبت فى عمليات الصدق. كذلك نجد أن صدق التكوين يثير طرق بحث جديدة فى جمع البيانات الصادقة ، وبالرغم من ذلك نجد ان الأساليب المستخدمة فى التحقق من صدق التكوين مألوفة وتقليدية ، وأن المجال الفعال ممتد ليشمل مختلف الإجراءات.

إن صدق تكوين الاختبار هو المدى الذى يوضح لنا ما يمكن أن يقدمه الاختبار من قياس للتكوين النظرى أو السمة. وأمثلة للتكوينات ، الاستعداد المدرسى ، الفهم الميكانيكى ، الطلاقة اللفظية ، والعصابية والقلق ، وأى تكوين واضح أن يفسر ويلاحظ الاستجابات بطريقة متسقة ، وهذا كله تشأ من توطيد العلاقات المتبادلة بين مقاييس السلوك. كذلك فإن صدق التكوين يتطلب معلومات متراكمة من مختلف المصادر. وأى بيانات واضحة ومطروحة عن بحث طبيعة السمة والعوامل المؤثرة فيها والمظاهر توضح المؤشر المناسب لذلك الصدق.

التغيرات النمائية:

إن المحك الرئيسى المستخدم فى صدق عدد من اختبارات الذكاء التقليدية هو محك (الفروق فى العمر) ومن هذه الاختبارات ، ستانفورد . بنية ومعظم اختبارات ما قبل المدرسة والتى تضبط العمر الزمنى لتحديد ما إذا كامت الدرجات تصف زيادة متقدمة فى الذكاء مع تقدم السن. حيث أنه إذا أوضحت نتائج درجات الاختبار زيادة فى الذكاء (أو فى القدرات المتوقع زيادتها) مع تقدم العمر خلال مرحلة الطفولة فإن ذلك يدل على صدق الاختبار. ونجد أن المفهوم الحقيقى لمقياس الأعمار الزمنية بناء على ما ذكرة بينية أنه افتراض أساسى لزيادة الذكاء مع العمر ، منذ الصغر وحتى الرشد.

وبالطبع ، نجد أن محك الفروق فى العمر ليس مناسب مع بعض الوظائف التى لا تظهر بوضوح مع تغيرات العمر. فعلى سبيل المثال ، فى نطاق قياس الشخصية ، نجد أن اختلاف العمر ضرورى ولكنة ليس شرط كافى للصدق. هكذا فإذا كانت درجات الاختبار تهمل التحسن مع العمر فمن المحتمل أن تبين أن الاختبار ليس صادق فى قياس القدرات التى صمم لاختبارها.

وعلى صعيد آخر ، لإثبات ان الاختبار يقيس بعض الأشياء التى تزيد مع العمر ، فإن هذا لا يحدد بدقة المساحة الحقيقية أو المجال الذى يغطية الاختبار ذلك لأن قياس الطول والوزن يصف أيضا الأتساق فى زيادة العمر ، بالرغم من أن ذلك لا يظهر بوضوح فى اختبارات الذكاء.

اخر نقطة تؤكد عليها متعلقة بتفسير محك العمر. أن الاختبار النفسى الصادق كمحك يقيس خصائص السلوك التى تزيد مع العمر تحت الشروط اللازمة للحياة فى أى نوع من البيئات والتى يقيسها الاختبار. ونتيجة لأن مختلف الثقافات تحفز النمو المتباين فى خصائص السلوك ، ذلك لأنه لا يمكن أن نفرض أن محك الفروق فى العمر واحد وشامل مثل جميع المحكات الآخرى. فهذا المحك مقيد بيئة الفرد الثقافية التى نيشق منها.

هناك أيضا تحليلات نمائية تستخدم كقاعدة فى صدق تكوين المقاييس المتدرجة (موضح فى الفصل الـ 3 و الـ 9) والافتراض الأساسى لكل المقاييس يتمثل فى إحراز مراحل تالية فى مفهوم التطور وهو شرط اساسى لإكتساب المصادرات المفاهيمية. وذلك يمثل سلسلة حقيقية فى محتوى هذه المقاييس. أن التكوين الصادق للمقاييس الرتيبة لابد أن تتضمن بيانات تجريبية عن النتيجة الثابتة فى خطوات متتابعة. وهذا يستلزم مقارنة أداء الأطفال فى مستويات مختلفة فى تطوير مفهوم الاختبار ، كبقاء أو صياغة الأهداف. هل الأطفال الذين يظهروا براعة فى مستوى متقدم ، يستطيعوا ان يظهروا تفوق فى المستويات المنخفضة.

الإرتباطات بين الاختبارات وبعضها:

إن الاتباطات يبين اختبار جديد واختبارات آخرى مشابهة احيانا يتم ذكرها. كدليل على أن الاختبار الجديد يقيس تقريبا نفس المجال الكلى من

السلوك كباقى الاختبارات المستخدمة بنفس الاسم مثل "اختبارات الذكاء أو اختبارات الاستعداد الميكانيكى" بخلاف ذلك نجد أن الارتباطات الموجودة فى صدق المحك التنبؤى هى ارتباطات متوسطة الارتفاع وليست مرتفعة جدا. وإذا كان الاختبار الجديد يرتبط مع اختبار آخر متاح ارتباط عالى جدا ، فإذا ذلك يوضح أن الاختبار الجديد يعتبر نسخة غير ضرورية.

إن الارتباطات مع الاختبارات الأخرى مستخدمة فى توضيح أن الاختبار الجديد غير متصل بعوامل أخرى خارجية تؤثر عليه. مثل اختبار الإستعداد الخاص أو اختبارات الشخصية لا يمكن أن يكون هناك ارتباط عالى بينهم وبين اختبارات الذكاء العام أو الأستعداد المدرسى. وبالمثل (الفهم القرائى) لا يجب أن يكون له تأثير على الاختبارات الأخرى ، وهكذا نجد أن الارتباطات مع اختبارات الذكاء العام ، القراءة ، الفهم الهجائى أحيانا تكون غير مباشرة أو ذات علاقة سالبة للصدق. وفى هذه الحالات الارتباطات العالية ممكن أن تجعل الاختبار مشكوك فيه. مع ذلك الارتباطات المنخفضة لا يمكن أن يؤكد على الصدق.

التحليل العاملى:

التحليلى العاملى هام ووثيق الصلة بإجراءات صدق التكوين ، وهو يوضح لنا معنى تماثل السمات النفسية. نجد أن التحليل العاملى هو أسلوب إحصائى دقيق لتحليل العلاقات المتبادلة لبيانات السلوك (المعلومات الخاصة بالسلوك) مثلا إذا كان لدينا (20) اختبار قد تم تطبيقهم على (300) شخص ، أو خطوة لحساب الارتباطات بين كل اختبار والاختبارات الأخرى ، هى فحص نتائج الجدول ، وهو يضم (190) علاقة ارتباط تكشف لنا عدد من الارتباطات لمجموعة معينة من الاختبارات. كذلك فهو اقتراح لمعرفة السمات المشتركة بين الاختبارات ، لذلك نجد أنه إذا كانت كل الاختبارات ، كأختبار المفردات

واختبار المتشابهات وتكملة الجملة ، بينهم ارتباطات عالية مع بعضهم البعض وارتباطات منخفضة مع جميع الاختبارات. ويمكننا أن نستنتج حضور عامل الفهم اللفظى. ذلك لأن تحليل جدول الارتباطات صعب وغير محدد ، فإنا نستخدم أكثر الأساليب الإحصائية التى ظهرت لتحديد العوامل المسئولة عن الارتباطات وهى أساليب التحليل العاملى والتى تم استخدامها فى بحث طبيعة الذكاء.

فى عمليات التحليل العاملى نجد عدد من المتغيرات أو الفئات تظهر فى مصطلحات يمكن أن يتم وصفها من خلال عدد من الاختبارات الأصلية (المفترضة) لعدد صغير من العوامل أو السمات المشتركة. فى المثال السابق ، نجد أنه يصبح لدينا 5 ، 6 عوامل ، تم حسابهم عن طريق العلاقات الإرتباطية بين 20 اختبار ، وكل شخص لابد أن نصف درجته فى الـ 5 ، 6 عوامل ، ويصبح لدينا (20) درجة أصلية. والسبب الرئيسى لهذا التحليل العاملى هو تبسيط وصف السلوك عن طريق تقليل عدد الفئات من خلال تبسيط عدد الاختبارات المتعددة إلى عدد قليل من العوامل أو السمات.

ونجد أنه بعد مطابقة العوامل ، يمكن أن تفيدنا فى وصف تركيب الاختبار ، فكل اختبار يمكن أن يميز العامل الرئيسى المحدد لدرجاته وذلك بواسطة درجة التشبع بالعامل وارتباط الاختبار مع أى عامل وهذا الارتباط أحيانا ما يوضح الصدق العاملى للاختبار المفردات (0.66) فإن الصدق العاملى يعنى أن اختبار المفردات صادق فى أنه يقيس سمة الفهم اللفظى بمقدار (0.66) وذلك يبين لنا أن الصدق العاملى اساسى فى ارتباط اختبار مع آخر. فهو يبين لنا اشتراك مجموعة من الاختبارات أو مؤشرات السلوك. وعن طريقها يمكن تحديد مجموعة من المتغيرات وتحليلها ، وتتضمن كلا من الاختبار وبيانات الاختبار الآخر ونجد أن تقديرات وقياس المحك يمكن استخدامه بكثرة عن طريق

اختبارات آخرى تبين الإتفاق العاملى للاختبار المستقل وإيجاد السمات المشتركة المقاسة.

الأتساق الداخلى:

فى الوصف المعلن لبعض الاختبارات ، خاصة اختبارات الشخصية ، قد تم توضيح أن الاختبار يمكن أن يكون صادق عن طريق طريقة الإتساق الداخلى. حيث أن الصفة المميزة لهذه الطريقة هى أن المحك لا يمكن أن يبين الدرجة الكلية للاختبار. وأحيانا ، نجد أن هناك تكيف فى استخدام طريقة المجموعة المضادة ، والمجموعتين الطرفتين التى تم اختبارها فى الأساس من خلال درجة الاختبار الكلية. ونجد أن الأداء المرتفع للجماعة المعيارية على مفردة الاختبار يظهر من خلال الجماعة المعيارية المنخفضة. ونجد ان المفردات المهملة تصف مستوى دلالة كبير لنسبة النجاح والذى يفوق الجماعة المعيارية المنخفضة. ونلاحظ أن الاجراءات المرتبطة يمن ايضا ان تستعمل من خلال ذلك الغرض. على سبيل المثال ، الارتباط بين (النجاح ـ الفشل) فى كل مفردة والدرجة الكلية للاختبار يمكن أن نحسبها. ومفردات الاختبار التى يتم اختبارها بهذه الطريقة يمكن أن تصق لنا الاتساق الداخلى، ذلك لأن كل مفردة تميز بين المستجيبين فى نفس الاتجاه.

استعمال آخر لمحك الاتساق الداخلى تتضح من خلال ارتباط درجات الاختبار الفرعى مع الدرجة الكلية. مثلا ، عدد من اختبارات الذكاء تتكون من مجموعة من الاختبارات الفرعية المختلفة. (مثل المفردات ، الحساب ، إكتمال الصور...) نقوم بتجميع درجاتهم كى نحصل على الدرجة الكلية للاختبار. وفى تكوين هذه الاختبارات ، نجد أن الدرجات فى كل اختبار فرعى غالبا ما تكون مرتبطة مع الدرجة الكلية. وكذلك فإن اى اختبار فرعى مرتبط مع الدرجة

الكلية بدرجة منخفضةنتجاهله أما إرتباطات الاختبارات الفرعية المتبقية مع الدرجة الكلية فهى تبين لنا الاتساق الداخلى كوسيلة صحيحة.

وذل يظهر لنا إرتباطات الاتساق الداخلى التى تعتمد على المفردات أو الاختبارات الفرعية ، تمثل فى الأساس مقاييس للتجانس. ذلك لأنها تساعدنا على وصف السلوك أو السمات فى الاختبار ، وكذلك درجة تجانس الاختبار التى تبين صدق التكوين. وبالرغم من ذلك ، نجد أن إسهام بيانات الاتساق الداخلى فى صدق الاختبار محدود. وفى حالة إنعدام البيانات الداخلية عن الاختبار ، نجد انه لا نستطيع أن نعلم ألا القليل عن ما يقيه الاختبار.

الصدق التقاربى والتميزى:

من خلال التحليل المتعمق لصدق التكوين ، قد أوضح D. T Campbell 1960 انه كى نبين صدق التكوين لابد أن لا نصف الارتباطات العالية للاختبار مع متغيرات آخرى فحسب ولكن أيضا ليس إرتباط دال مع متغيرات مختلفة. وفى مقال لـ D. T Campbell and Fiske (1959) قد قاموا بوصف العمليات المكونة للصدق التقاربى والصدق التميزى. فمثلا ارتباط اختبار التفكير الكمى الدرجات فى الحساب يوضح لنا الصدق التقاربى. وفى نفس الاختبار ، نجد أن الصدق التمييزى يظهر منخفض بين درجات الاختبار ودرجات اختبار الفهم القرائى. ذلك لأن القدرة اللغوية ليس لها علاقة بالاختبار الذى صمم لقياس التفكير الكمى.

وذلك يوضح لنا ان الشرط الأساسى للإرتباط المنخفض مع المتغيرات الغير متصلة بالاختبار هو بحث الارتباطات بواسطة الإجراءات الإضافية المستخدمة فى صدق المحتوى. أيضا الصدق التمييزى يظهر بوضوح من خلال

اتصال بصدق اختبارات الشخصية والتى بها بعض المتغيرات الغير متصلة بالاختبار والتى يمكن أن تؤثر على الدرجات فى الأساليب المتعددة.

اقترح كلا من Campbell and Fiske عام (1959) تصميم تجريبى نظامى كمدخل مزدوج وثنائى للصدق التقاربى والتمييزى ، وكذلك فقط أطلقوا عليه (مصفوفة الطرق والسمات المتعددة) ونجد أن هذا الإجراء يتطلب تقييم اثنين او اكثر من السمات عن طريق اثنين او اكثر من الطرق. وفى مثال افتراضى زودنا به Campbell and Fiske يفيد فى توضيح الإجراءات. جدول (5 - 1) يصف كل العلاقات الممكنة التى حصلنا عليها عن طريق درجات 3 سمات للضخية مثل (A) السيطرة أو الهمنية (B) الإجتماعية (C) دافعية الإنجاز. أما الـ 3 طرق فهى (1) بيان مفصل عن تقرير الذات (2) الأسلوب الأسقاطى (3) تقييم الرفاق. ونلاحظ أن السمة (A1) تشير إلى درجات السيطرة فى بيان عن تقرير الذات أما السمة (A2) فهى تبين لنا سمة السيطرة المقاسة بالاختبار الأسقاطى أما (C3) فهى توضح لنا تقييم الرفاق لسمة دافعية الإنجاز.

هذه الارتباطات الافتراضية الموجودة فى الجدول (5 - 1) تتضمن معاملات ثبات (وهى موضوعة بين قوسين ، فى موازاة الخط القطرى الرئيسى) ومعاملات الصدق (وهى مطبوعة بأحرف سوداء إلى الأمام من اقصر ثلاثة خطوط مائلة) الجدول أيضا يتضمن أرتباطات بين مختلف السمات المقاسة بواسطة مختلف الطرق (كما فى المثلثات المتقطعة) ولأجل الحصول على صدق التكوين الفرضى ، نجد أن معاملات الصدق لابد أن توضح ارتباطات عالية بين مختلف السمات المقاسة مختلف الطرق ، كذلك يجب أيضا أن تكون الارتباطات اكثر ارتفاعا بين مختلف السمات المقاسة بنفس الطريقة (البيان التفصيلى عن تقرير الذات) ودرجات سمة الهيمنة عن طريق (الاختبار الأسقاطى) لابد ان نعطى ارتباط مرتفع. أكثر من الارتباط بين درجات الهيمنة والاجتماعية من خلال

البيان التفصيلى عن تقرير الذات. إذا صور الارتباط الأخير اختلاف الطرق الشائعة والتى عبرت عن ارتباطات مرتفعة. فإن ذلك يمكن أن يوضح لنا كما المثال ، أن درجات الفرد فى هذا البيان المفصل يتأثر على نحو ملائم ببعض العوامل العاملة التى ليس لها علاقة بالموضوع مثل القدرة على فهم الأسئلة أو المرغوبية الاجتماعية ومحاولة الفرد لإظهار التأييد الواضح فى كافة السمات.

	السمات Traits	Method 1 A_1	Method 1 B_1	Method 1 C_1	Method 2 A_2	Method 2 B_2	Method 2 C_2	Method 3 A_3	Method 3 B_3	Method 3 C_3
Method 1	A_1	(.89)								
	B_1	.51	(.89)							
	C_1	.38	.37	(.76)						
Method 2	A_2	.57	.22	.09	(.93)					
	B_2	.22	.57	.10	.68	(.94)				
	C_2	.11	.11	.46	.59	.58	(.84)			
Method 3	A_3	.56	.22	.11	.67	.42	.33	(.94)		
	B_3	.23	.58	.12	.43	.66	.34	.67	(.92)	
	C_3	.11	.11	.45	.34	.32	.58	.58	.60	(.85)

التدخل التجريبى:

يوجد مصدر إضافى لبيانات صدق التكوين وهو يظهر لنا من خلال التجارب عن طريق تأثير بعض المتغيرات المنتقاه على درجات الاختبار. وفى التحقق من صدق الاختبار نستخدم (برنامج للتعليم الفردى) على سبيل المثال ، يوجد مدخل يمكن من خلاله مقارنة درجات الاختبار القبلى مع درجات الاختبار

البعدى. ونجد أن هذا الأساس المنطقى يظهر من خلال انخفاض الدرجات فى الاختبار القبلى والأداء قبل إعطاء التعليمات. والدرجات العالية فى الاختبار البعدى. هذه العلاقة يمكن أن تختبر مفردة واحدة فى الاختبار. ونجد أن نسبة كبيرة من الممتحنين يفشلوا فى مفردة فى الاختبار القبلى ويتفوقوا فى الاختبار البعدى. أما إذا فشل الأفراد بشكل عمومى على المفردات فى كلا الاختبارين فتكون هذه المفردات صعبة جدا (ذو مستوى صعب) وإذا تفوقا فى كلا الاختبارين فهذا يعنى أنهما أكثر سهولة فى قياسهما للغرض من الاختبار. وإذا كانت نسبة الممتحنين تدل على تفوق فى الأداء على المفردة فى الاختبار القبلى ، وتدل على الفشل فى نفس المفردة فى الاختبار البعدى ، فإن ذلك يرجع أحيانا لوجود خطأ أما بواسطة المفردات أو التعليمات ، أو الأثنين معا.

وفى اختبار صمم لقياس (الميل للقلق) يمكن تطبيقه على شخص يتعرض بشكل متتالى لموقف صمم لاستثارة القلق ، كتطبيق الاختبار تحت شروط التشويش والإنفعال وحالات آخرى. فنجد أن درجات الاختبار المبدئى يمكن أن ترتبط ببعض العوامل الفسيولوجية للقلق أثناء وبعد الاختبار. والفرض المتعلق باختبار القلق يمكن تقييمة بتطبيقة قبل وبعد خبرة استثارة القلق ، إذا ارتفعت درجات الاختبار ارتفاع ذو دلالة فى إعادة الاختبار. فإن النتائج الايجابية تدل على أن درجة الاختبار تعكس مستوى القلق. ومن خلال التجرية لابد أن تبين أن درجات الاختبار تعكس المستوى الحالى للقلق ، وبطريقة آخرى ، التجارب يمكنها أن تصمم لاختبار الفروض الآخرى فيما يتعلق بالسمة المقاسة بواسطة الاختبار.

النمذجة بالمعادلات البنائية:

بالإضافة إلى تحديد بنية الاختبار والأداء على المحك ، شئ هام جدا لابد أن نأخذه فى إعتبارنا وهو العلاقات بين المكونات والمسارة الذى يؤثر فى أداء المحك

J. P Campbell, 1990a, Messick, 1989, Schmidt, Hunter &)
Outerbridge 1986) على سبيل المثال ، اهتمام الشخص فى منطقة معينة يمكن أن يؤثر على الأداء الوظيفى من خلال زيادة تعلمه معلومات حقيقية متصلة بالمجال أو من خلال اكتساب مهارات إجرائية متطلبة ، أو من خلال الدافعية اللازمة لبذل الجهد الأقصى للقيام بالأنشطة الضرورية. لدراسة كيفية تأثير السمات الشخصية التى تؤدى للأداء الجيد أو الأداء الضعيف يساهم فى فهم سبب ارتفاع أو انخفاض صدق الاختبار فى مواقف معينة. وهذا التحليل يكون من السهل بإستخدام أجراء احصائى يعرف باسم (النمذجة بالمعادلة البنائية) والتى بدأت فى الظهور من الثمانينات (1980 . 1990) وكما ذكرنا أن هذا الإجراء مرتبط بتحليل المسار وكلاهما مصممات للدراسة "كنماذج سببية".

ما هى الشئ الذى صمم لإنجازة اجراء (نمذجة المعادلات البنائية) كيف تم إنشاؤه؟ نجد انه فى الإحصاءات الأولية ، يوجد شئ واحد يوضح لنا أن الارتباطات ليست سببية فقط اى (لا تشير إلى السببة فقط) ويظهر لنا ذلك من خلال مثال مألوف عن ارتباطات الأعمار الغير منطقية. ما بين الأطفال من المدرسة من سن 6 . 14 سنة والاتباط بين الجودة والارتفاع فى مادة الحساب والتى تتسم بالارتفاع بالرغم من عدم القدرة على انها تأثير المتغيرات الأخرى. والنتائج النهائية لارتباط الأعمار ، قد أظهرت الارتباط مع مقدرا التعليم. وفى محاولة لفحص العلاقات السببية فى الأبحاث. قام الباحثين فى عام 1960 - 1970 باستخدام (التصميم التجريبى المتقاطع) (D. T Campbell, Stanby
1966, Cook & Campbell, 1976 P. P. 284 - 293) على سبيل المثال ، لتحليل تأثير الأسباب المتبادلة ما بين استعداد الطلاب تجاه الحساب وأدائهم الحسابى. هذه المقاييس (الاستعداد نحو الحساب والأداء فى الحساب) يمكن أن نحصل عليهم فى نقطتين فى نفس الوقت. ذلك لأنه يمكن إيجاده من خلال

(الإرتباط المتقاطع) بين الاستعداء الحسابى فى الوقت الأول (1) والأداء الحسابى فى الوقت (2) وبين الأداء الحسابى فى الوقت الأول (1) والاستعداء الحسابى فى الوقت (2) ونجد أن حجم هذه الارتباطات تشير إلى شدة التأثير فى كل الاتجاهات. وخلال عدد قليل من السنوات ، قد ظهرت طريقة لتقدير تأثير متغيرين على بعضهم البعض.

نجد أنه قبل ذلك كشف كلا من التحليل الإحصائى والمنطقى عن أوجه قصور خطيرة فى استخدام طريقة الارتباطات المتقاطعة. بالرغم من أن هذه الطريقة صممت بعناية إلا أننا نجد أن نظام الاتباطات يؤدى بنا إلى نتائج غير سليمة ويؤدى بنا أيضا إلى وجود إجراءات خاطئة خاصة بالعلاقات السببية .(Rogosa 1980)

ومن ضمن مصادر الخطأ فى هذا الإجراء ، أنه أولا ، يهمل الارتباط بين المتغيرات البدائية والمتغيرات النهائية. ثانيا ، يهمل ثبات المتغيرات بإختلاف الزمن. ثالثا يهمل دراسة بعض المتغيرات الغير مقاسة والموثرة على المتغيرات التى يقيسها ، مثل العمر ومقدار التعليم فى المثال الذى سبق ذكرة ، إن نمذجة المعادلاة البنائية تمدنا بطرق آخرى يمكنها أن تتجنب هذه الصعوبات. حيث أن هذا الإجراء فى الأساس يعتمد على استخدام معادلات الانحدار فى التنبؤ بالمتغيرات المستقلة فى (التصميم التجريبى المتقاطع أو النماذج السببية الآخرى) هذا الأسلوب يعتمد على معادلات الإنحدار المتعددة فى إيجاد معاملات الإنحدار. حيث أنه بتلك الوسيلة تندمج كل الارتباطات بين المتغيرات ، كذلك فإن كلا من الأخطاء البسيطة والمقاسة يمكن أن تظهر من خلال التفسير. هذا ونجد أن نقص الاحتياطات المسبقة يمن أن يؤدى إلى زيادة المتغيرات السببية الغير مقاسة Bentler, 1988, L.R. James et al., 1982 Loehlin, 1992,) .(Rogosa, 1979

أو خطوة فى النمذجة بالمعادلات البنائية هى تصميم نموذج افتراضى يضم العلاقات السببية التى سوف نختبرها. وذلك هام لأن ذلك النموذج يعتمد على الألفة الكاملة من خلال معرفتنا بمتغيرات ومواقف البحث. ونجد أن العلاقات الإفتراضية لابد أن تكون فى ضوء أساس نظرى وخلفية نظرية معينة. الاختبار الحقيقي لهذا النموذج يتم من خلال حل مجموعة من معادلات الإنحدار المتأنية. حيث أن النماذج السببية، عادتا ما تكون مليئة بالمعادلات الغير معروفة والتى تتيح الفرصة إلى التوصل للحلول عن طريق نماذج اختبارية خاصة. والنموذج يقارن بين مصفوفة الارتباط الأصلية (المفترضة) والتجريبية للوصول إلى معلومات تخص حسن المطابقة (Mac Callum, Wegener, Uchind & Fabrigar 1993) وهذه النماذج الإحصائية تكافئ نماذج توضح اختلاف المسارات السببية وتمدنا بالتفسيرات البديلة عن تأثيرات الملاحظة التجريبية. وبخصوص البحث عن المعرفة الخاصة بالمكان، لابد للفرد أن يقيم النماذج البديلة ويتم ذلك عن طريق قبولة لها وكذلك المغزى الحقيقى منها.

وهناك مميزة آخرى للنمذجة بالمعادلات البنائية وهى أنه يمكن حساب العلاقات السببية إما بين التكوينات أو بين المتغيرات المقاسة. على سبيل المثال، كى نقوم بتقييم إستعداد الطلاب تجاه مادة الحساب، فإننا استخدام عدد من المقاييس، كمقاييس للاهتمامات، ومفهوم الذات كمقياس للاستعداد الحسابى، وبعض تأثيرات المتغيرات المتصلة بالموضوع المقاس. ونلاحظ ان الاشتراك بين هذه المقاييس يمكن أن نجدها من خلال تكوين الاستعداد الفردى تجاه مادة الحساب، والذى يتصل بالأداء الحسابى التالى. وهذا الاستخدام للتكوينات يمدنا بتقديرات موثوق بها وثابتة والتى بدورها توضح الخطأ والاختلافات الخاصة بالمقاييس.

ونجد ان النمذجة بالمعادلات البنائية تمثل إجراء له مستقبل باهر من خلال (الاتحاد النظرى، التجارب) والمداخل الإحصائية التى يعتمد عليها

كذلك فهو يتجه على نحو واسع إلى التطبيق فى المشكلات التى تنتمى لمجالات كثيرة وذلك لتطوير علم النفس الاجتماعى والصناعى والشخصية وكذلك التعليمى. وهذه المحاولات تتطلب إتحاد وإيضاح وتيسير للإجراءات المستخدمة، وهى فى الحقيقة فى طريقها لتحقيق ذلك (E. G., Joreskog & Sorbom 1993).

الإسهامات فى علم النفس المعرفى:

لقد شهد عام 1970 تقارب كبير بين علم النفس التجريبى والقياس النفسى، وذلك بدأ من خلال الإسهامات الخاصة بفهم وتقييم تكوين اختبارات الذكاء، الاستعدادات الخاصة (Ronning, Glover, Comoley, & Wott, 1987, R. E. Snow & Loh - Mon, 1989) وفى عام 1950، بدأ علماء النفس المعرفيين فى استعمال بعض المفاهيم الخاصة بالعمليات المعرفية. وذلك لدراسة حل المشكلات ونجد أن بعض الأبحاث صممت كبرامج لمقارنة تنفيذ هذه العمليات. وفيما يتصل بالتفكير الإنساني الزائف، فيمكن للبرامج أن تكشف عن الأداء الكاذب (الزائف) للأفراد فى مستويات مختلفة من المهارة وعن طريق هذه البرامج يمكننا أن نتنبأ بعدد الأخطاء التى نقوم بها وأنواعها وكذلك الوقت المستغرق فى الاستجابات المختلفة. وفى تصميم هذا البرنامج نجد أن الباحث عادتا ما يبدأ بتحليل المهمة والتى يمكن أن تتضمن استبطان طريقة الفرد فى حل المشكلات. وكذلك تفكير الفرد بصوت مرتفع عند تناول الموضوعات واستخدامه للإجراءات الدقيقة فى الملاحظة. وعن طريق مقارنة الأداء عند كل الأطفال والشباب أو الخبراء والمبتدئين فى حل المشكلات (المسائل) يمكننا اختبار الفروض وبحثها عن طريق معرفة كيفية قيام الأفراد بهذه المهام. مثال لهذه المهام التى يتم بحثها عن طريق هذه الطرق، المهام التى تتضمن ألغاز محيرة، مشاكل فى المنطق، ألعاب الشطرنج، مسائل فى الجبر، مسائل فى الفيزياء، مشكلات

تشخصية فى ممارسة الطب (مسائل متعلقة بالتشخيص الطبى) (Chi, Glaser, Farr, 1988, J. H. Larkin, Mc Dermoott, Simon, & Simon, 1980a, 1980b, Newell & Simon, 1972, Simon, 1976).

إن تطابق المتغيرات بواسطة هذه الأبحاث يتضمن عمليات خاصة بالمهارات الإجرائية والمعرفة التقريرية (الحقائق والمعلومات) ونجد أن النماذج المعرفية الخاصة بالعمليات العقلية المستخدمة فى إنجاز المهامة يمكن أن يوضح الطريقة ، العمليات المنظمة ، مستودع المعرفة ، وكذلك كيفية تمثيل هذه المعرفة فى الذاكرة ، وكيفية استرجاعها عند الحاجة؟ كذلك فإن زيادة الإنتباه يوضح لنا كيفية إستدعاء العمليات المنفذة والتى تدل على تحكم الفرد فى اختبار العمليات ، التصورات (التمثيلات) والاستراتيجيات الخاصة بالمهمة.

وفى عام 1970 بدا بعض علماء النفس المعرفى فى تطبيق (تحليلات المهامة) والأساليب المتعلقة بالحاسب فى استكشاف ما تقيسه اختبارات الذكاء (See Resnick, 1976, Sternberg, 1981, 1984, 1985b) ونجد أن نتائج هذه البحوث قد أسهمت تدريجيا فى تقدير مناسب لتكوين الاختبار واستخدامه.

أن تضمنيات البحث المعرفى المتعلقة بصدق التكوين قد تم وصفها وشرحها بصفة خاصة من خلال ما قام به (Embreston 1983, 1986, 1995a) ومن الملاحظ انه قد قام بتحديد التعليمات الخاصة بصدق التكوين بشكل كامل فى مدخلة ، والذى اقترح فيه 2 من **المظاهر المبدئية لصدق الاختبار:**

1- تمثيل التكوين.

2- مدى مجموعة المعارف المتصلة بالتكوين.

ونجد أن تعليمات الصدق المرتبط بالتكوين ترتكز على مدى مجموعة المعارف المتصلة به. والتى تتعلق بعلاقات الأداء على الاختبار داخل (شبكة مجموعة المعارف المتصلة) بالمتغيرات الآخرى. كالعلاقات التى يتم بحثها بشكل عام من خلال إرتباطات درجات الاختبار مع المقاييس الاخرى ، والتى تتضمن الأداء على المحك والبيانات الحقيقية الآخرى.

على صعيد آخر ، نجد أن موضوع تمثيل التكوين ، نحتاجه فى مطابقة عمليات معرفية معينة ومقادير معرفية مستخدمة فى إنجاز المهام من خلال مفردات الاختبار. ونلاحظ أن تحليل المهام يمكن ان يتم تجريبيا وذلك لاستنتاج كل التحليلات. وأمثلة للإجراءات التى يمكن أن تشتمل عليها معالجة المهام المعقدة ، تقديم المهام الجزئية ، أو توفير أسباب لتبديل متطلبات المهامة ، النماذج الحسابية التى تم تطويرها لقياس تنوع مكونات الإستجابة نحو أداء المفردة. ويستخدم هذا الإجراء فى تحليل المهامة المعرفية عن طريق (تحليل البروتوكولات) (Ericsson, 1987, Ericsson & Simon, 1993, Van Someren, Barnard & Sandberg, 1994).

فى هذا الإجراء ، يتعلم الأفراد أسلوب (التفكير بصوت مرتفع) كى يتمكنوا من إنجاز المهامة أو حل المسائل. وهذه المهامات يمكن تقديرها من خلال عمليات عقلية كالمضاعفة (ضرب الأعداد) تذكر اجزاء من أحداث ماضية ، أو اكتشاف أسباب القصور فى تجهيز الإجابة الصحيحة للإسئلة. والنتيجة الثانية لهذا الإجراء هو إمكانية إكتشاف أن نفس مفردات الاختبار يمكن أن تستدعى عمليات معرفية مختلفة تماما للمستجيين ، وذلك نتيجة لوجود تنوع فى الخلفيات التجريبية للأفراد.

ما الذى نستنتجة حول إسهام علم النفس المعرفى حتى هذه النقطة فى الصدق المرتبط بالتكوين؟ حتى ذلك الحين مازال فى المرحلة التمهيدية ، مداخل

للعمليات المعرفية يمكن أن تستنتج مفاهيم تساعدنا على توجية البحث. كذلك فنجد أن المعرفيين ركزوا إنتباههم ويشكل واضح على عمليات الاستجابة والتى تختلف عن أبحاث علماء القياس النفسى الذين أوضحوا تركيز التعليم على الناتج النهائى للتفكير. إن تحليل أداء الاختبار فى تعبيرات خاصة بالعمليات المعرفية لابد ان يقوى ويبسط فهمنا عن كيفية قياس الاختبار. مع ذلك نجد أن تحليل أداء الأفراد فى المستوى المبدئى لابد ان يؤدى فى النهاية لإمكانية التحديد الدقيق لدرجات الفرد فى مواطن الضعف والقوة والعلاج التشخيصى لاستخدام الاختبارات (Embretson, 1987, 1994, Estes, 1974, Pelle - Grion & Glaser, 1979, Sternberg & Weil, 1980).

وتباعا لذلك ، نجد أنه لابد من تسهيل برامج التدريب كى تلائم حاجات الأفراد ، ونجد أن العلاقة بين القياس النفسى والمداخل المعرفية يمكن أن تصور لنا أولا ، من خلال وجهة النظر المعتمدة على تطبيق البحث ، وتطبيق ذلك عمليا فى تكامل. وهنا نجد أن الكل يمثل مدخل ممتاز وذلك من خلال التقديرات المختلفة للأهداف. ثانيا ، من خلال وجهة النظر المعتمدة على النظرية والبحث والتى تبين لنا أن العلاقة يمكن أن تبدو علاقة عكسية ، وهنا الكل يساعد فى التفسير وفى التوضيح والاتحاد وذلك كله يزيد من قيمة فهمنا التصرف المتعلق بالذكاء.

التكامل والتداخل

مقارنة إجراءات الصدق:

لقد قمنا بدراسة طرق مختلفة للصدق وذلك يجعلنا نطرح السؤال التالى ، "كيف يكون صدق هذا الاختبار" وكى نوضح المظاهر المختلفة والمميزة لإجراءات الصدق المختلفة ، دعنا نطبق ذلك عمليا من خلال محتوى اختبارية (50) مسألة حسابية. يمكننا أن نستخدم 4 طرق فى هذا الاختبار ، عن طريق تحديد نوع إجراءات الصدق الملائمة لكل طريقة منهم ، وذلك موضح فى جدول (5 - 2) وهذا المثال يركز إنتباهنا على حقيقية اساسية ، وهى أن اختبار إجراء الصدق يعتمد على استخدامنا لدرجات الاختبار.

ونلاحظ أن نفس الاختبار ، عندما يستخدم لغايات مختلفة ، لابد أن يكون صادق فى مختلف الطرق. فإذا كان الاختبار التحصيلى المستخدم فى التنبؤ بالأداء اللاحق فى مستوى تعليم عالى ، كاختبار طلاب المدرسة الثانوية كى يلتحقوا بالجامعة ، فإننا نحتاج ذلك أما للتقييم فى ضوء محك الأاء الجامعى اللاحق أو بلغة آخرى صدق المحتوى.

شمولية التكوين الصادق:

المثال الموضح فى جدول (5 - 2) يركز على الأختلافات بين أنواع الإجراءات الصادقة. علاوة على ذلك ، نجد أن بحث هذه الإجراءات يصف لنا أن (صدق المحتوى وصدق المحك والصدق المرتبط بالتكوين) ليست متطابقة فى كل الأنواع ولكنها منطقية أو متميزة. وعلى العكس ، نجد ان الصدق المرتبط بالتكوين يعبر عن مفهوم السمولية ، وذلك لأنه يتضمن كافة الأنواع الأخرى وكل الأساليب المحددة من خلال تحليل المحتوى أو من خلال إرتباطات المحك المستخدم فى القياس ، يمكن ان يميل مرة آخرى لمطابقة التكوين. على سبيل المثال

، إرتباطات اختبار الاستعداد الميكانيكى مع الأداء فى الطرق المختلفة وفى تنوع الوظائف تشارك فى فهم التكوين الخاص بالاختبار. إن مطابقة التكوين يمكن أن يتم من خلال مقارنة درجات المجموعة المتضادة (عمال ناجحين . عمال غير ناجين).

جدول (5 - 2)

تحقيق صدق إختبار حساب فردى من خلال أهداف مختلفة

إثبات الصدق	أسئلة توضيحية	هدف إجراء الاختبار
- وصف المحتوى.	- ما مقدار التعليم الذى حصل عليه ديك فى الماضى؟	- اختبار تحصيلى فى المدرسة الإبتدائية (حساب).
- التنبؤ بالمحك (زمنى).	- ماذا سيتعلم جاك فى المستقبل؟	- اختبار استعداد للتنبؤ بمستوى الرياضيات فى المدرسة الثانوية
- التنبؤ بمحك (تلازمى).	- هل أداء بيل يدل على صعوبات خاصة؟	- أسلوب لتشخيص صعوبات التعليم.
	- كيف تكون درجة هيلين المتصلة بمؤشرات القدرة على التفكير؟	- مقياس للتفكير الكمى.
- تطابق التكوين.		

ونلاحظ أن الصدق فى ضوء المحكات العملية المختلفة عادتا ما يكون موضح فى تعليمات الاختبار، وذلك كى يساعد مستخدمى الاختبار على فهم ما يقيسه الاختبار. بالرغم من أن الفرد لا يمكنة بشكل مباشر أن يهتم بالتنبؤ بأى المحكات المستخدمة، فى الدراسة التى يمكنة الإعتماد عليها فى توضيح مفهوم السلوك الخاص الذى يقيسه الاختبار. إلا اننا إذا قمنا بتتبع هذه الفكرة، فإننا نستطيع ان نرى أن كل استخدامات الاختبار وكل التفسيرات المتعلقة

بدرجات الاختبار متضمنة فى صدق التكوين (J. P. Campbell, 1990a, (Guion, 1991, Messick, 1980b, 1988, 1989, Tenopyr, 1986 لذلك نجد أن المعنى التفسيرى لدرجات الاختبار غالبا ما يعتمد على التكوينات. والتى تتنوع فى سعة أفق التفكير والقدرة على التعميم من خلال ملاحظة السلوك الخاص والبيئات المحيطة.د

ولقد ناقش (Messick (1989, 1980b مصطلح الصدق وذلك كى يدل على تفسير الهدف من الاختبار. ونجد أن ذلك يساعدنا فى تحديد صدق التكوين. كذلك فإن الإجراءات الأخرى المتعلقة بمصطلح الصدق يمكن أن يتم تجميعها بشكل تقليدى ، وقد اكد Messick انه يمكننا عن طريق وصف العبارات بصفة خاصة أن تشير إلى أن صدق المحتوى يعبر عن (محتوى متصل بالموضوع ، محتوى يغطى الموضوع) وذلك يشير على التوالى إلى التخصيص والتمثيل ، أما الصدق المرتبط بالمحك يمكن ان يعبر عن (فائدة التنبؤ ، فائدة التشخيص) وذلك يتوافق مع الصدق التنبؤى والصدق التلازمى. من ناحية آخرى نجد أن ، الاختلافات بين مختلف إجراءات الصدق تساعدنا فى التعرف على الهدف من الاختبارات. وهى تكون موضحة فى التعليمات الخاصة بكل اختبار.

ومن زاوية آخرى ، نجد أن الاستعمال أو التطبيق العملى المباشر يركز على وصف المحتوى (كما فى التقدير التعليمى) أو التنبؤ بالمحك (كما فى الاختبار الوظيفى) ونلاحظ أن استخدام التكوينات فى عرض غرض معين أو استخدامه يعتبر اكثر تأثيرا لقياس أداء الاختبار الخاص. كذلك فإن بحث المحك يوضح لنا أن كلا من أجراءات المحك ودرجات الاختبار يمكن أن تؤثر بشكل واضح فى اختبار التكوينات علاوة على ذلك ، نجد الاستقصاء عن العلاقات السببية بين التكوينات ، كما النمذجة بالمعادلات البنائية ، يمكن أن تؤدى بنا إلى فهم ومعرفة مهمة أو نتيجة الاختبارات.د

صدق العمليات المكونة للإختبار:

نلاحظ أن صدق الاختبار يتطلب عند تطويره إجراءات متعددة ، والتى يمكن أن تستخدم بشكل تتابعى فى مختلف المراحل المكونة للاختبار (Anastasi, 1986, Guion, 1991, Jackson, 1973, N. G. Peterson et al., 1990) وكما نعلم أن الصدق ينشأ فى الاختبار أما من خلال محك خارجى أو من خلال تحديده فى المراحل النهائية لتطوير الاختبار. ونلاحظ أن عمليات الصدق تبدأ من خلال الصياغة المفصلة للسمة أو تعريف التكوين والذى ينشق من خلال نظرية نفسية معينة ، أو بحث قبلى ، أو حتى ملاحظات منظمة وتحليل للسلوك الخاص. وتكون مفردات الاختبار مجهزة للعرض كتكوين بعدما نقوم بتحليلها عن طريق اختبار بعض المؤثرات الخارجية أو صدق المفردات من خلال مجموعة مفردات أولية يتم تحليلها عن طريق أساليب إحصائية ، وآخر شئ هو تحديد صدق الدرجات وتفسير توافق الدرجات فى ضوء محك خارجى.

ونجد أن معظم المعلومات المجمعة فى عمليات تطوير واستخدام الاختبار متصلة بصدق هذا الاختبار. وبالطبع ، نجد أن البيانات عن الأتساق الداخلى وعن ثبات الاختبار تساعدنا فى إيجاد تجانس التكوين وثباته الموقت. وكذلك فإن المعايير يمكن أن تمدنا بمواصفات إضافية للتكوين ، خاصة إذا تضمنت تمييز بين البيانات المعيارية للجماعات الفرعية التى تم تصنيفها على أساس (العمر أو النوع) أو أى متغيرات آخرى تؤثر على التاريخ التجريبى للأفراد. وبتلك الوسيلة نستنتج أداء الفرد على الاختبار. ونجد أن تفسير الدرجة التى يحصل عليها الفرد تحتاج إلى تراكم الملاحظات الفردية والبحوث.

النتائج الفردية والمجتمعية لإجراء الاختبار:

إن بعض علماء القياس النفسى قد وجهو إنتباهنا بأن هناك صورة أو ميزة متضمنة فى صدق الاختبار وهى يطلق عليها نتائج إجراء الاختبار على الفرد وعلى المجتمع. ومن الجدير بالذكر، أنهم شددوا على هذا العرض لمفهوم الصدق، خاصة العالم Messick (1980b, 1988, 1989, 1995) والذى أكد على الوضع الغير مقصود للنتائج فى الاستخدام الخاص للاختبارات والتى يمكن أن تكون غير مناسبة للأفراد أو لعدد من الجماعات الأخرى نتيجة لإختلاف خلفيتهم الإختبارية. وقد إتضح لنا فى تقرير لجنة الخبراء المجتمعين بواسطة (NRC) والتى قامت بتقدير طلبات الوظائف ودراسة ذلك الوضع بشكل كامل وأكثر شمولية.

ومن ناحية آخرى، نجد أن التضمينات الإجتماعية والعرقية للاختبار المستخدم من غير ريب تحتاج إلى مؤسسات إجتماعية واسعة الانتشار وبعض المظاهر المتعلقة بهذه المسائل (موضحة فى الفصل الـ 18) ويظهر لنا فى (الفصل الـ 6) من خلال (تحيز الاختبار) وكى ندخل كل هذه التأثيرات فى مفهوم الصدق، سنجد أنه سيظهر لنا ان الأسئلة المتعلقة بهذه المسألة، لا يمكن الإجابة عليها بواسطة الأساليب الإحصائية فقد. وهكذا نجد أن الحكم على أساس توازن الأهداف يتم عن طريق تناول طرق لمجالجة نظم القيمة (Mullen & Roth, 1991, Zeichmeister & Johnson, 1992) كل الطرق المتطلبة للحكم على الإنسان من خلال، المناقشة المنظمة، التسوية على نحو كاف، حل النزاعات، وفى الأساس نجد أن اختلاف قيم النظم قد تم تمثيلها على نحو كافى، عند ضم الإجراءات التجريبية والإحصائية المحتملة لتحديد الصدق عن طريق تقييم النتائج الاجتماعية والعرقية للاختبار المستقل والتى تستخدم حل يتسم بالتشويش والأرباك.

فى النهاية ، نجد أنه من خلال دراسة الصعوبات والمسائل الهامة يمكننا تحديد الدور الرئيسى للاختبار المستعمل (موضحة فى الفصل 1) وعندما تكون الأحكام على القيم متطلبة ، خاصة فى الحالات الفردية التى تعتمد على الاختبار المستعمل. نجد أنه فى حالة إختيار الاختبارات وترجمة النتائج ، يمكننا أن نتحكم فى الأساليب الإجرائية للأختبار المستعمل وكذلك يمكننا عرض قيمة التوجية وحساسية المجتمعية لأى اختبار مستعمل يمكن أن يساعد بشكل دال فى الموافقة على استخدام الاختبارات وذلك يحدث ليس من خلال وجهة نظر علمية فقط وإنما من خلال وجهة النظر العرقية أيضا .

الفصل الثامن

الاختبارات الجماعية

يتم تطبيق الاختبارات الفردية مثل ستانفورد بينيه ومقياس وكسلر فى العيادة نفسها بينما الاختبارات الجماعية تستخدم أساسا فى النظام التعليمى ، الصناعة ، الخدمات الحكومية ، الخدمات العسكرية. فى أثناء الحرب العالمية الأولى تم بناء اختبار شامل يشمل جيش الفا وجيش بيتا للاستخدام فى جيش الولايات المتحدة. جيش ألفا هو اختبار لفظىصمم من أجل أغراض الحماية وتحديد المكان الملائم لكل شخص. جيش بيتا هو اختبار غير لفظى يستخدم مع الأشخاص الذين لا يستطيعون الإجابة على اختبار ألفا بسبب خلفيتهم فى اللغة الأجنبية أو أن يكونوا أميين. لقد تم إنشاء نموذج بواسطة هذه الاختبارات وإتبع فى التمنية المتتالية عدد كبير من الاختبارات الجماعية التى تستخدم فى التطبيق المدنى.

إن الأشكال المتفق عليها فى القانون المدنى استمرت فى الاستخدام عدة عصور فى اختبارات الجيش الأصلية وفى خدمات الجيش العسكرية. وقد تم استخدام اختبار التأهيل للقوات المسلحة (AFQT) وقد نما بطريقة متتالية مثل الداة التمهيدية التى تبعت بواسطة بطارية تقييم الاستعدادات المتعددة من أجل مهمة فى التخصصات المهنية. اختبار AFQT زود بمقياس واحد قائم على عدد متساوى من المعانى ، الحساب ، العلاقات المكانية ، مفردات للقدرات المهنية. بطارية الاستعداد المهنى للقوات المسلحة ASVAB تستخدم فى كل خدمات الجيش كبطارية تصنيفية. بعض الاختبارات الفرعية لـ ASVAB تستخدم كمكون تأهيلى مشترك فى التصنيف الشخصى كل خدمة يقوم بها الفرد تنتمى وتندمج مع الاختبارات الفرعية لتلائم احتياجاتها الخاصة المهنية المحددة.

في هذا الفصل يجب أولا أن ندرس الاختلافات الأساسية بين الاختبارات الفردية والجماعية. هذا سوف يتم بواسطة بحث الإجراءات التى نشأت من أجل اختبار جماعى للأفراد وقد يتم بواسطة استخدام الحاسوب فى برنامج الاختبار. يجب أن نقدم بعض الأمثلة الواضحة للاختبارات الجماعية واسعة الرتبة من أجل الاستخدام العام. فى النهاية ، نحن يجب أن ندرس الاتجاه الذى يهتم بتنمية الاختبارات واستخدام الاختبارات التى تلائم الجماعة أى نستخدم الاختبارات التى تميز بين الأفراد فى الاختبارات الجماعية كما فى الاختبارات الفردية (التى نوقشت فى الفصل 8) لذلك يتم الاهتمام بدمج الاختبارات المصممة كقياسات شاملة لقدرة استيعابية واحدة مثل (الذكاء . الاستعداد الدراسى) ببطارية الاستعدادات المتعددة. اختبارات الذكاء تتميز بالمرونة حيث أن أداة واحدة يمكن أن تحدد درجات هذه الاختبارات بعرض متنوع من العام إلى الخاص لتناسب التنوع الواسع لأغراض الاختبار والموقف الذى يوجد فيه.

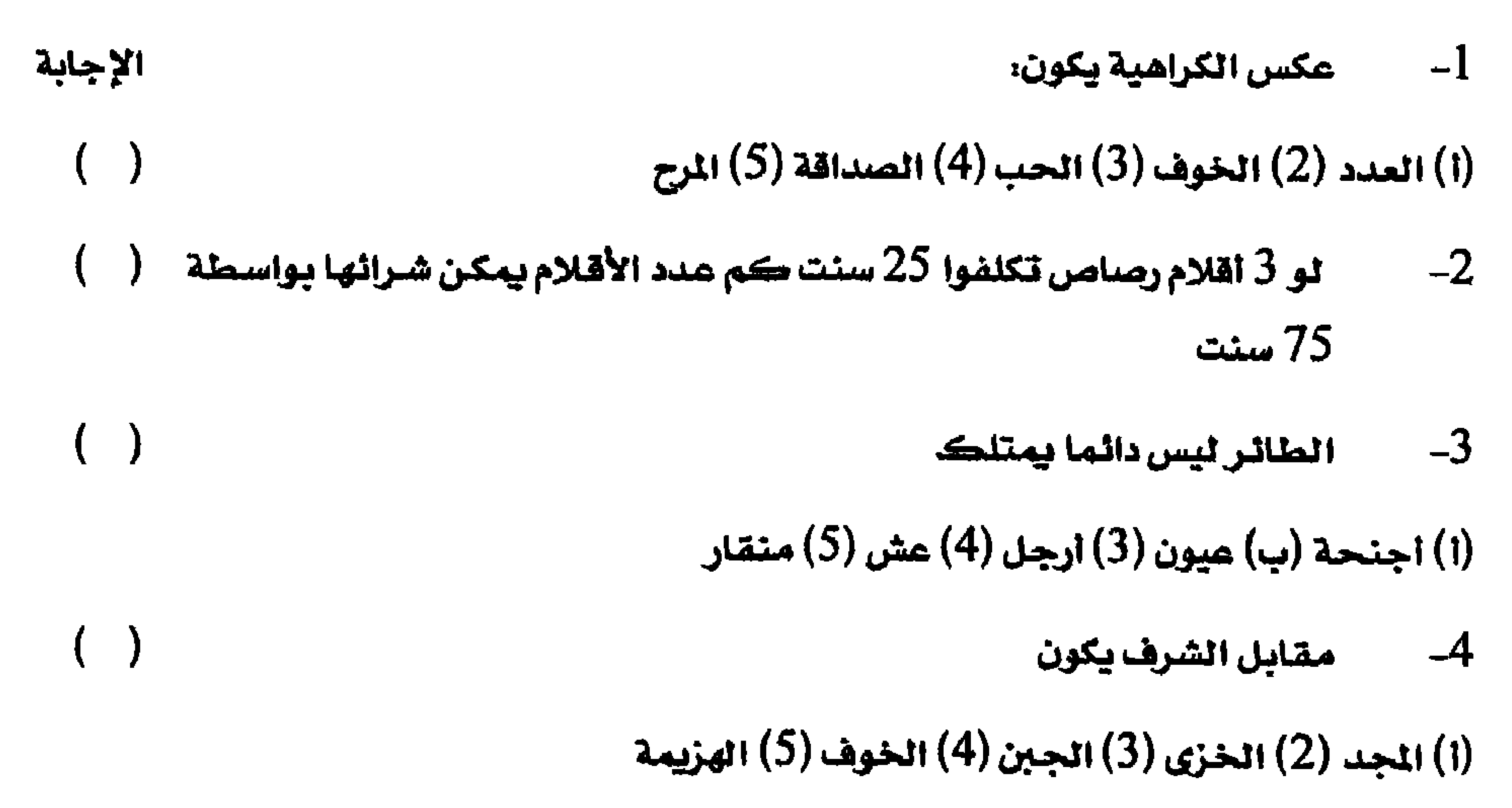
الاختبارات الجماعية مقابل الاختبارات الفردية

الاختلافات النموذجية فى تصميم الاختبار:

الاختبارات الجماعية تختلف عن الاختبارات الفردية فى شكل وترتيب البنود. لقد استخدمت الأسئلة التى تتلقى إجابات حرة فى الاختبارات الجماعية الأولى ، اليوم ، الاختبارات الجماعية النموذجية تستخدم الدرجات. هناك اختلاف أخر كبير بين الاختبارات الجماعية والفردية هو التحكم فى صعوبة المفردات (البنود). فى الاختبارات الفردية المفحوص يتبع القواعد المهمة وغير المهمة. يقوم الفاحص بفحص البنود التى تناسب قدرات المفحوص. فى الاختبارات الجماعية يتم ترتيب بنود المحتوى ترتيب متزايد فى الصعوبة خلال الاختبارات الفرعية المنفصله والمعتمدة على الزمن. لابد من الاهتمام بتنظيم مفردات الاختبار مثل:

اختبار المعانى ، الحساب والاختبارات المكانية ويتم الاهتمام بإكمال البنود السهلة ثم التوجه إلى البنود الصعبة حتى نستطيع أن نتعامل بطريقة جيدة مع الوقت

هناك صعوبة عملةي قد واجهت الاختبارات الفرعية المنفصله هى أن المفحوصين الأقل حرصا والأقل خبرة ربما يصنعوا أخطاء الوقت هذه الأخطاء قد تحدث بكثرة وقد تكون نسبية. لكى يتم استخدام زمن محدد للاختبار مع ترتيب الاختيار بطريقة جيدة من السهولة إلى الصعوبة يقوم مؤسسى الاختبار باستخدام الشكل الحلزونى الشامل واحد من الاختبارت الأولية التى تقدم هذا التصميم كان اختبار أوتيس (Otis) لإدارة الذات وهو اختبار للقدرت العقلية العامة والذى من أسمه نجد أنه يقلل من دور الفاحص إلى أقل حد. يبدأ الشكل الحلزونى الشامل من أسهل البنود ثم يتبع النوع الأسهل التالى وهكذا حتى يصل إلى مستوى حلزونى من الصعوبة كما يشرح بأسفل.



1- عكس الكراهية يكون: الإجابة

(ا) العدد (2) الخوف (3) الحب (4) الصداقة (5) المرح ( )

2- لو 3 اقلام رصاص تكلفوا 25 سنت كم عدد الأقلام يمكن شرائها بواسطة 75 سنت ( )

3- الطائر ليس دائما يمتلك ( )

(ا) اجنحة (ب) عيون (3) ارجل (4) عش (5) منقار

4- مقابل الشرف يكون ( )

(ا) المجد (2) الخزى (3) الجبن (4) الخوف (5) الهزيمة

لكى نتجنب التعليمات المكررة فى كل مفردة ولنقلل من التغيرات التى تحدث فى مجموعة التعليمات فبعض الاختبارات تستخدم ترتيب حلزونى شامل ليس للمفرة الواحدة ولكن إلى تجمعات من (5- 10) مفردات.

مميزات الاختبارات الجماعية:

الاختبارات الجماعية صممت كأدوات للاختبار الشامل بالمقارنة بالاختبارات الفردية وكلا منهم يمتلك مزايا وعيوب على الجانب الإيجابى نجد أن الاختبارات الجماعية تتم فى وقت واحد على العديد من الأشخاص. برامج اختبار المقياس الكبير صنعت بواسطة تنمية أساليب الاختبارات الجماعية. باستخدام المفردات المطبوعة والاستجابات البسيطة التى يتم تسجيلها على ورقة الإجابة أو على الكمبيوتر أو على كتب الاختبار يتم إلغاء العلاقة بين الفاحص والمفحوص (علاقة واحد إلى واحد).

الطريقة الثانية التى فيها الاختبارات الجماعية جعلت الاختبار الشامل يتميز بالسهولة كانت من خلال تبسيط دور الفاحص على العكس من التدريب المكثف والخبرة المطلوبة لمقياس ستانفورد بينيه على سبيل المثال، معظم الاختبارات الجماعية تتطلب القدرة على قراءة التعليمات البسيطة إلى المفحوصين ثم المحافظة على دقة الوقت. بعض جلسات التدريب التمهيدى يكون مرغوب فيها لأن المفحوصين الذين لا يمتلكون خبرة فى تطبيق الاختبارات قد ينحرفوا من خلال إهمال بعض المستويات التى ربما تؤثر على نتائج الاختبار. على الرغم من أن دور الفاحص يكون قليل لكنه الاختبار الجماعى يمكنه أن يزود بحالات متماثلة كثيرة من الاختبار الفردى. عند استخدام التعليمات المسجلة على شريط وإدارة الكمبيوتر فهذا يعطى فرصة لتحديد المعايير التى يتم استخدامها كما إنه يقلل من تغير الفاحص فى اختبار ذات المقياس الكبير كما أن درجات الاختبار تتم بطريقة نموذجية أكثر.

موضوعية فى الاختبار الجماعى:

من زاوية أخرى الاختبارات الجماعية تزودنا بنموذج جيد أكثر من الاختبارات الفردية لأن الاختبارات الجماعية تتميز بالسهولة النسبية والسرعة فى جمع البيانات. من المألوف أن نختبر العينات الممثله فى عملية التقنين. فى معظم الاختبارات الجماعية المقننة الحديثة أحيانا العينات المقننة تتراوح بين 10.000 ، 200.000 بالمقارنة مع 1.000 إلى 8.000 حالة متراكمة عند تقنين مقاييس الذكاء الفردية.

عيوب الاختبارات الجماعية:

بالرغم من أن هذه الاختبارات تمتلك سمات عديدة مرغوب فيها ولها وظائف كثيرة لا يمكن الاستغناء عنها لكن نجد أن الفاحص فى الاختبارات الجماعية يمتلك فرصة قليلة فى إنشاء علاقات بينه وبين المفحوص ولا يحدث تعاون بينهم ولا يحافظ على اهتمامات وميول المفحوص لذلك أى حالة مؤقته يتعرض لها المفحوص مثل المرض والأوهام والضيق ربما تتدخل فى أداء الفرد على الاختبار والتى توجد بنسبة كبيرة فى الاختبارات الجماعية عن الاختبارات الفردية. الأشخاص الغير معتادين على أداء الاختبارات الفردية. لقد ظهرت بعض الأدلة المقترحة التى تؤكد على أن الأطفال التغير متزنين عاطفيا ربما يكون أدائهم أفضل فى الاختبار الفردى أكثر من الاختبار الجماعى (Bower 1969; Willis 1970).

لقد تم مهاجمة الاختبارات الجماعية بسبب القيود المفروضة على استجابات المفحوصين ولقد تم أيضا توجيه النقد بطريقة مباشرة ومحددة إلى بنود الاختيار المتعدد إلى أنواع المفردات المقاسة مثل المتشابهات والمترادفات والتصنيفات (Hoffman 1962; Lafave 1966).

إذا حدث خلاف فإنه يكون مثل المفردات التى تعوق المفكر الأصلى الذى يرى تداعيات فريدة فى الإجابات. يجب أن تذكر بطريقة ثانوية أنه لو حدث هذا يجب ان يكون حدوثه نادرا من وجهة نظر تحليل المفردات وصدق البيانات. لو حدث هذا فى مفردة او اثنين فقد يظهر تأثيره على المقدار الكلى للمفحوص وذلك فى الاختبار الفردى. بعض النقد ركز على أهمية تحليل الأخطاء وتساءل عن الأسباب لماذا الفرد اختار هذه الإجابة كما فى مدخل Piagetian النموذجى (Sigel, 1963).

مما لا شك فيه أن الاختبارات الجماعية تزود بقليل من الفرص من أجل الملاحظات المباشرة لسلوك المفحوصين ومن أجل تحديد أسباب الأداء غير النموذجى. من أجل هذه الأسباب والأسباب الأخرى ، عندما نضع القرارات المهمة عن الأفراد فإنه يتم توفير الاختبارات الجماعية إما مع فحص اختبار فردى للحالات المشكوك فيها أو مع معلومات إضافية من مصادر أخرى.

مازال هناك قصور آخر فى الاختبارات الجماعية التقليدية وهو أن الاختبار يكون ناقصا للمرنة فكل مفحوص يتم اختباره فى كل المفردات ولكن إذا ركز المفحوص على المفردات التى تناب قدراته فيصبح وقت الاختبار له تأثير كبير وله أهمية فمثلا فى المفردات السهلة يتجنب المفحوص الملل من العمل الذى يقوم به ولكن إذا كانت المفردات تتسم بالصعوبة أى يتم المحاولة والتفكير فيها فإن المفحوص قد يشعر بالقلق والإحباط من هذه المفردات. لكن الاختبار الفردى الذى يتم بطريقة نموذجية يزود الفاحص بالقدرة على اختيار المفردات المناسبة والتى تناسب استجابات الطلاب على الاختبارات السابقة. هذا التميزي بين الاختبارات الفردية والجماعية هام عند تصميم الاختبار الذى يهتم بقدرات الإنسان على نطاق واسع.

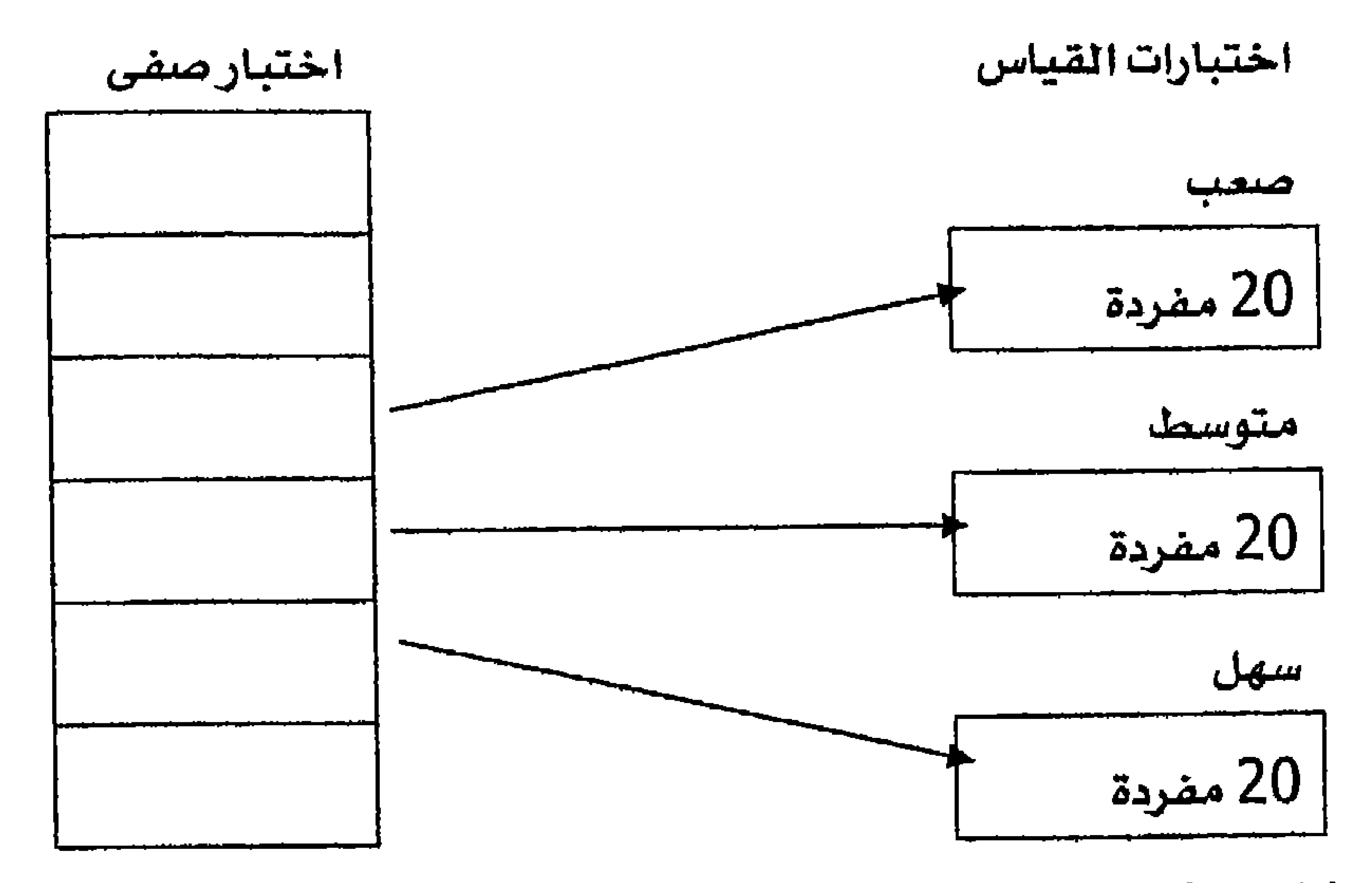
الإدار المعتمدة على الكمبيوتر واختبار التكيف

الاختبارات المتصله بشكل فردي:

لقد تم بذل مجهود كبير لدمج بعض مميزات الاختبار الجماعى مع الاختبار الفردى ولقد تم اكتساب العديد من الأساليب طبقا لذلك. ظهر اهتمام كبير بالمفردات التى تركز على خصائص استجابات المفحوصين فى اختبار الفردى. هذا المدخل لقب باختبار تكيفى ، تتابعى ، منفرع ، متصل على شكل ، فردى ، مبرمج ، حركى ، أو اختبار مشروط بالإجابة من الممكن أن تصمم الاختبارات الجماعية التحريرية (بالورقة والقلم) والتى تدمج هذه الإجراءات التكيفية هذه الأساليب تستخدم لتقدير إدارة الاختبار بالكمبيوتر (Cleary, Linn & Rock 1968; Lord 1971a).

الاختبار التكيفى يعتمد على تنوع واسع للنماذج الإجرائية (Dewitt & Weiss, 1974; Larkin & Weiss, 1974; Weiss, 1974; Weiss & Betz 1973) مثال بسيط يشمل اختبار المرحلتين يتم شرحه فى الشكل (Betz 1973) (10- 1) فى هذا الاختبار يأخذ كل مفحوص 10 مفردات من الاختبار كل مفحوص يوجه إلى واحد من ثلاث اختبارات القياس ذات العشرين مفردة على مستويات عديدة من الصعوبة لذلك كل شخص يأخذ 30 مفردة بالرغم من أن الاختبار الكلى يشمل 70 مفردة بترتيب مختلف يتم توضيحه فى الاختبار الهرمى المبين فى الشكل (10- 2) فى هذه الحالة كل المفحوصين يبدأوا بالمفردات ذات الصعوبة المتوسطة فإذا كانت إجابة المفحوص صحيحة فإنه يتجه إلى الموضوع التالى الأكثر صعوبة ولكن إذا فشل المفحوص فى الإجابة فإنه يتجه إلى الموضوع الأكثر سهولة التالى. هذا الإجراء يكرر بعد كل إجابة مفردة حتى يحصل الفرد على 10 درجات. الشكل يبين اختبر المرحلة اعاشرة الذى يحصل فيه

كل مفحص على 10 مفردات من 55 مفردة فى الاختبار. فى الشكل (10- 2) الخط الثقيل يوضح الطريقة المتبع بواسطة المفحوص ويتم وضع إجابته فى قائمة (+ ، -) على طول المقدمة



شكل 10- 1) اختبار تكيفى ذو مرحلتين مع ثلاث مستويات قياس كل مفحوص يأخذ واحد من الاختبار الصفى مع اختبار قياسى

شكل (10- 2) نموذج الاختبار الهرمى الخط الثقيل يبينه طريقة أخذ الاختبار إجابات المفردات بجدول عبر القمة

الاختبار التكيفى المحسوب (CAT)

إن المتغيرات العددية لكل من نماذج الاختبار التكيفى المشروحه فى الشكل (10- 1) (10- 2) استخدم معها أشكال الورقة والقلم وايضا الكمبيوتر لكن النماذج المركبة لا تستخدم النماذج المنشأة سابقا فيتم استخدام الاختبار التكيفى معها (Embretson, 1992; B.F. Green 1993; Wainer *et al,* 1990) هذه الإجراءات الخاصة بالاختبار التكيفى المحسوب تستخدم أساليب نظرية البند والاستجابة (IRT) تهتم بتجميع مجموعة من المفردات من أجل اختبار الأفراد ومن أجل تقدير أداء الأفراد. من أجل كل مفردة فى المجموعة يتم تقدير قدرة الفرد المطلوبة من أجل (50- 50) فرصة لاجتياز هذه المفرداتز تقدر هذه القدرة من خلال قياس مستقبلات الفرد لاجتياز هذه المفردات. الذى يعكس

لنا مستوى صعوبة الاختبار هو وجود قيم التمييز ووجود درجة من التخمين مرتبطة بهذه المفرداتز المعلومات المتاحة عن المفردات تشير إلى دقة القياس والمعلموات الدالة على الاختبار تكون مجموعة معلومات عن المفردات. المعلومات المتاحة عن المفردات تشير إلى دقة القياس والمعلومات الدالة على الاختبار تكون مجموعة معلومات عن المفردات التى تخدم نفس الغرض الذى يقيسه الاختبار مثل خطأ المستوى التقليدى للقياس (SEM). يتم متابعة إجابة كل مفردة كما أن الكمبيوتر يختار المفردات التالية على أساس الإجابة الكلية كما أن الكمبيوتر يختار المفردات التالية على أساس الإجابة الكلية للفرد والاختبار يستمر فى إضافة مفردات جديدة حيث نجد أن المعلومات الدالة على الاختبار تصل إلى مستوى جيد من الدقة لذلك كل الأفراد يتم قياسهم على نفس مستوى الدقة.

لا يتم تقدير الفرد على أساس عدد المفردات التى أجابها بطريقة صحيحة ولكن على أساس مستوى صعوبة هذه المفردات وأيضا الخصائص السيكومترية لهذه المفردات. أن درجات الاختبار الكلى تكون مشتقة من تقييم قدرة الفرد المقابلة لكل مفردة تم اجتيازها. هذه الدرجات يتم مقارنتها بكل الأشخاص الذين يتم فحصهم فى مجموعة المفردات بعيدا عن المفردات المحددة لكل فرد. الإجراءا تالفعالة لإنشاء اداة CAT يتم تسهيلها من خلال استخدام العديد من برامج الحاسوب المتاحة مثل MICRO CAT بواسطة مجلس الأنظمة التقيمية.

إن البحث بواسطة طرق متنوعة يشير إلى أن الاختبار التكيفى الفردى يمكنه أن يحقق نفس الثبات والصدق كاختبارات مهنية مع اقل عدد من المفردات وفى اقل وقت للاختبار. كما أن هذه الطرق تمدنا بدقة القياس من اجل الأفراد حيث توضح اقصى ارتفاع واقصى انخفاض لقدرات الأفراد التى يتم توضيحها بواسطة الاختبار (Lord 1970, 971a, 1971b, 1971c; Weiss, 1982)

يوجد بحث اخر يؤخذ فى الاعتبار هو الذى يوضح الترابط بين اشكال CAT المنشأة جيدا للاختبار ومن هذه الأشكال الورقة والقلم "تحريرية" كما أن هذه الترابطات تكون عالية مثل معاملات الثبات لمعظم الاختبارات إن نتائج البحث تقترح أن هذه الأشكال يتم استخدامها مع مختلف الاختبارات (Mead & Drasgow, 1993) لكن هناك بعض مواقف الاختبار الـ CAT لا يكون مناسب معها مثل اختبارات السرعة التى توضح تنوع الأشخاص (Wainer 1993b) لقد ظهر اهتمام خاص بتقييم ادوات CAT (Green, Bock, Humphreys, Linn & Reckase 1984).

الاختبار التكيفى يستخدم بطريقة خاصة مع البرامج التعليمية الفردية التى يتميز فيها الطلاب بالسرعة وهذا يتطلب أن تكون مفردات الاختبار على مستويات مختلفة من الصعوبة. الاختبارات التى تتم بواسطة الكمبيوتر تجعل من الممكن أن نوقف الاختبار بمجرد أن إجابات الطالب تعطى معلومات كافية لإصدار قرار بإتقان الفرد للمحتوى هناك استخدامات أخرى للاختبار التكيفى المدار بالكمبيوتر أنه يهتم بالبحث عن النشط المناسب للفرد ويتم تنمية ذه الأنشطة فى العديد من الإطارات. بواسطة الاختبار التكيفى المحوسب تم تطبيق اختبار تعليمى على طلاب الجامعة من اجل أن نحدد دخول طلاب السنة الأولى بالجامعة إلى دورات انجليزية ورياضية مناسبة (Smittle 1990; Ward Kline & Flaugher, 1986) بسبب انتقاء مفردات فردية فهذا الاختيار يتطلب قليل من الوقت ويزودنا بدرجة الفرد فى الحال لذلك يتم وضع الطلاب فى دورات او اقسام أفضل تناسب مستوى قدراتهم أو أدائهم الحالى.

يوجد تطبيق اخر مهم CAT من خلال برامج التصنيف التى تستخدم فى انتقاء الأشخاص المناسبين فى مجالات الصناعة والحكومة والخدمات العسكرية.

إن CAT تكون مناسبة من أجل 3 أسباب على الأقل:

1- التدفق الثابت للطلاب المتقدمين للاختبار لذلك تعوق الاختبارات التى تتم على مجموعات كبيرة فى أى وقت وفى أى مكان.

2- يتم التعامل مع معدلات واسعة من مستويات القدرات.

3- زيادة أمن الاختبار لأن كل متقدم يستقبل مجموعة مختلفة من صيغ المفردات من المجموعات الكبيرة المخزونة فى الكمبيوتر.

إن البحث التفسيرى نحو تنمية نسخة CAT لبطارية الاستعداد المهنى لخدمات الجيش ASVAB يكون فى تقدم تدريجى (McBirde & Matin 1983; Moreno, Wetzel, McBride & Weiss 1984; Wiskoff & Schratz 1989) لقد تطورت الـ CAT من أجل الاختبارات الجماعية مثل اختبارات الاستعدادات الفارقة التى وضعت فى الجزء الأخير من هذا الفصل والتىتستخدم من أجل التطبيقات العملية ومن أجل البحث عن طبيعة ومصادر الاختلافات الفردية بذلك نجد أن الـ CAT تتميز بمميزات عددية.

البطاريات متعددة المستويات

مقدمة:

على عكس معظم المقاييس الفردية والاختبارات التكيفية المدارة بالكمبويتر (المحسوب) فإن الاختبارات الجماعية التقليدية تمثل نفس المفردات لكل المفحوصين. بغض النظر عن إجابات الأفراد فن اى اختبار جماعى يجب أن يحتوى على مجال نسبى من الصعوبة وأن يكون مناسب لقدرات الأفراد. لكى تزود بقياسات مقارنة لتنمية القدرات العقلية على معدل واسع فقد تم إنشاء تسلسلات من البطاريات متعددة المستويات. لذلك اى فرد يختبر وفقا للمستوى المناسب له ولكن المستويات الأخرى يمكن أن تستخدم من أجل إعادة اختبار نفس الشخص فى السنوات التالية أو من أجل المقارنة بين مجموعات العمر المختلفة. الحقيقة أن البطارية الناجحة تزودنا بقعدة أساسية عن الأفراد فى مختلف أعمارهم ومختلف توزيعاتهم الرتبية. يجبأن نتعرف على أن الموائمة بين صعوبة المفردات وقدرة المفحوص تتم بواسطة البطاريات متعددة المستويات على عكس الإجراءات الفردية المتبعه فى CAT حيث نجد أن الموائمة تكون قائمة على معرفة سابقة عن المفحوص مثل عمره أو درجته المدرسية أكثر من استجاباتهم على الاختبار.

الباريات متعددة المسويات تكون مناسبة للاستخدام فى المدارس حيث يتم مقارنة درجات الأفراد على سنوات متعددة من أجل هذا السبب المستويات التى تتم بطريقة نموذجية يتم وصفها فى مصطلحات الرتب. معظم البطاريات متعددة المستويات تزودنا بدرجة معقولة من الارتباط مع الاهتمام بوظائف القدرات العقلية أو المحتوى. الدرجات يتم التعبير عنها فى عبارات نفس المقياس. يتم استخدام نظرية البند والاستداية IRT.

بطريقة متزايدة من أجل الحصول على درجات الأفراد والقدرة على مقارنته خلال مستوى معين. فى عملية تقنين الاختبار فإن مجموعة الطلاب تعطى مستويات متداخله للاختبار لتزودنا بالبيانات المهمة عن الأفراد. فالعينات المعيارية التى فحصت على مستويات متداخله تكون متكافئة للاختبارات المقاسه بطريقة مستقلة. فالمستويات الفردية تأخذ احيانا من 1 إلى 3 مستويات لكن فى البطارية متعددة المستويات قد تمتد من روضة الأطفال إلى دخول الكلية.

معظم البطاريات حصدت درجات لمقياس كلى مماثل لنسبة الذكاء التقليدية فى الاختبارات الفردية. بعض البطاريات تزودنا بأنواع متعددة من النماذج تشمل درجات مئوية متكافئة كدرجات للمقياس بالإضافة إلى الدرجة الكلية للاختبار. لذلك نجد أن معظم البطاريات تحصد درجات منفصلة لفظية وكمية أو لغوية وغير لغوية ونجد أن الأداء الفردى فى الاختبارات اللفظية وفى الأنواع الأخرى للاختبارات الفرعية ربما تكون غير متشابهة بالأخص فى المستويات العليا.

أسماء البطاريات تكون أيضا ذات اهتمام كبير مثل هذه المصطلحات الذكاء. القدرة العامة. القدرة العقلية. النضج العقلى. القدرة المدرسية كل هذه المصطلحات تستخدم لكى تحدد نوع الاختبار. فى القياس السيكومترى نجد أن هذه المصطلحات تكون فعليا مترادفة ومتغيرة. من الجدير بالانتباه أن معظم البطاريات الحديثة قامت باستبدال مصطلح الذكاء بألقاب محددة لكن التغير قد يوضح لنا أن مصطلح الذكاء قد اكتسب معانى عديدة ومتنوعة وقد يؤدى إلى التفسير الخاطئ لدرجات الاختبار. البطاريات متعددة المستويات تختبر المهارات العقلية الكبرى المطلوبة من أجل العمل المدرسى فوظيفتهم الأولية تقيم قدرة الفرد على القراءة فى المدرسة عند اى مرحلة فى العملية التعليمية.

البطاريات التمثيلية:

طبيعة بطاريات القدرات متعددة المستويات يتم شرحها مع 3 بطاريات وضعت فى الجدول (10- 1) هذه البطاريات اختبرت بسبب حداثة طبيعتها وارتفاع الإجراءات البنائية للاختبار وحجم تمثيل عينات التقنين مازال هناك صفة أخرى يجب أن تؤخذ فى الاعتبار هى أن هذه البطاريات يتم قياسها بطريقة متزامنة أى فى نفس الوقت مع واحد أو اثنين من البطاريات متعددة المستويات لاختبارات التحصيل التعليمى من أجل نفس الدرجات وبالتالى قد تنشأ مراسلة بين المجموعتين فى الدرجات ونجد أن الأداتين يمكن أن يستخدموا بطريقة مترابطة حتى يتم استكشاف التنمية التعليمية للطالب بشرط أن تؤثر فيه.

لقد تم البحث عن الثبات والصدق لهذه البطاريات بواسطة إجراءات مناسبة فمعاملات الثبات الخاصة بكيودر. ريشاردسون تحسب من خلال الدرجات الكلية والدرجات على اثنين أو ثلاث بطاريات ذات محتويات منفصلة. عندما تكون ترابطات إعادة الاختبار عالية فهذا يشير إلى ثبات مقبول. لكن صدق الاختبار يتم حسابه من خلال الترابط بين درجات المدرسة مع اختبارات التحصيل مما يشير إلى صدق تنبؤى جيد. فالترابط بين الدرجات الجزئية والتحليلات العاملية يكشف لنا عن عامل عام خلال كل بطارية كلية.

جدول (10- 1) البطاريات النموذجية ذات المستويات المتعددة

بطارية	المستوى	عدد المستويات	تكون معيار مشترك مع
اختبار القدرات المدرسية لـ اوتيس ولينون (OLSAT)	k-12	7	تسلسلات اختبار تحصيل ستانفورد
اختبار القدرات المعرفية (COGAT)	k-3 3-12	2 8	اختبارات Iowa للمهارات الأساسية (المستوى k-9) اختبارات للتحصيل والكفاءة (المستوى 9- 12) اختبارات Iowa للتنمية التعليمية (المستوى 9- 12)
اختبار المهارات المعرفية (TC S/2)	2- 12	6	اختبارات كاليفورنيا للتحصيل الاختبارات الشاملة للمهارات الأساسية

هناك أيضا اختبار أولى منفصل للمهارات المعرفية (PTCS) مع مجموعات مختلفة من الاختبارات التى تناسب روضة الأطفال ومستويات المرحلة الأولى.

محتوى الاختبار النموذجي عند مستويات مختلفة:

لقد تم استخدام الاختبارات الجماعية مع الأطفال ذات العمر الصغير مثل روضة الأطفال ومستويات المرحلة الأولى لكن الاختبار الفردى مطلوب لدى الأطفال فى عمر ما قبل المدرسة لكى تنشأ علاقة مع هؤلاء الأطفال ويتم التحدث معهم بطريقة شفوية بذلك يتم تحديد المفردات المناسبة لهؤلاء الأطفال زفى

عمر 5 او 6 سنوات من الممكن أن ننظم اختبارات مطبوعة لمجموعات صغيرة لا يزيد عددهم عن 10 او 15 طفل فى هذه الاختبارات الفاحص يوجه انتباه فردى إلى الأطفال ليتأكد من أن توجيهاتهم تكون متبعة وإذا كان من الضرورى أن يتم اختيار المجموعات الكبيرة فإنه من الممكن استخدام واحد أو اثنين كمساعدين للفاحص.

الاختبارات الجماعية الخاصة بالمستوى الأول تكون مناسبة لروضة الأطفال وأول درجتين أو ثلاث درجات من المدرسة الابتدائية وفى هذه الاختبارات يزود كل طفل بكتيب طبعت عليه صور ورسوم بيانية تشكل مفردات الاختبار وحتى يتأكد الفاحص أو المراقب من أن التعليمات قد فهمت بطريقة لائقة فإنه يراجع درجات الأفراد على بعض مفردات الاختبار. الأطفال يضعون علامة على الإجابة الصحيحة فى كتيب الاختبار باستخدام القلم الرصاص أو الكربون ويتم الاهتمام بالتناسق الحركى البسيط كما فى رسم الخطوط التى تربط بين نقطتين وبذلك نجد أن الاختبارات التى وضعت من أجل المستوى الأول لا تتطلب قراءة أو كتابة من المفحوص.

معظم بطاريات القدرات متعددة المستويات تشمل اختبارات مناسبة للمستوى الأول فأنواع مفردات الاختبار التى استخدمت فى هذا المستوى تشرح فى الشكل (10- 3) التى توضح بعض مفردات العينة من مستوى A لـ OLSAT (بطارية اختبار القدرات المدرسية لأوتيس ولينون) المناسبة لروضة الأطفال. لمعرفة نمو القدرات العقلية بطريقة سريعة والتى تحدث فى الأعمار الأولى فالطبعة الأخيرة لـ OLSAT تزودنا بأربعة مستويات منفصلة من (A.B.C.D.) من اجل روضة الأطفال ومن اجل الصف الدراسى 1 ، 2 ، 3 من المدرسة الابتدائية.

إن هذا يمثل اختلاف كبير بين الطبعات الأولى لهذه البطارية والبطاريات الأخرى ذات المستويات المتعددة. فى المستوى A كل التعليمات تعطى شفهيا

بواسطة الفاحص وعلى الطفل أن يجيب على الأسئلة من خلال ملئ الدائرة تحت صورة الإجابة الصحيحة كما يرى فى الشكل (10- 3).

لقد تم شرح 4 من 10 مفردات فى المستوى A فالاختبار الكلى يتطلب حوالى 75 دقيقة ومن الممكن إدارته على مرحلتين من خلال ان كلا منهما يأخذ 5 دقائق استراحة بعد أول 15- 20 دقيقة كما ان هناك اختبار ممارسة مع أنواع متشابهة من المفردات والتعليمات التى تعطى فى اليوم الأول. المفردات التى تختبر تعرض فى الشكل (10- 3) تكون مفردات بسيطة نسبيا وتستخدم لتجعل المفردات مألوفة بالنسبة للأطفال خاصة عند مواجهتهم للاختبار نفسه. التفسيرات التى أعطيت فى الشكل (10- 3) تكون مكدسة بطريقة كبيرة بالتعليمات الشفهية المفصلة ووصف واضح لمحتوى الاختبار لكن هناك اختلافات قليلة لتسهل فه موجذب انتباه الأطفال الصغار على سبيل المثال ، الصفحات والصفوف لا تحدد فقط بالأرقام لكن بواسطة صور دقيقة لأشياء مألوفة مل فنجان ، مقص أو حذاء ، وعلى كل طفل أن يضع علامة تحت الإجابة الصحيحة.

الاختبارات التى تستخدم فى المدرسة الابتدائية تمتد من الصف الثالث أو الرابع إلى أعلى وتمتلك كثير من الأشياء المشتركة فى كل من المحتوى والتصميم العام. نتيجة للجهل الوظيفي نجد أن الاختبارات التى تكون لفظية أساسا فى محتواها معظمها يشتمل على مشاكل حسابية أو اختبارات رقمية ، بالإضافة لذلك ، هناك بطاريات قليلة تزودنا باختبارات غير مقروءة صممت لتقييم القدرات الاستدلالية فى الأطفال الذين يمتلكون خلفية لغوية أجنبةي ، لديهم إعاقات فى القراءة أو إعاقات تعليمية أخرى.

إن المفردات التى تناسب مستوى الطلاب فى المدرسة الابتدائية شرحت فى الشكل (11- 4) هذه المفردات تكون نموذجية كاختبار القدرات المعرفية (COGAT) الذى يوجد فى المستويات المتوسطة كما فى الشكل (10- 1)

يشمل هذا الاختبار مستويين كل مستوى يتم طباعته ف ىكتيب منفصل فنجد أن المفحوصين يأخذون مستويات مختلفة وكل مفحوص يبدأ وينتهى بمجموعات مختلفة من المفردات. يتم اختبار المفحوصين بمفردات متوسطة الصعوبةحيث أن التمييز يكون أكثر تأثيرا. المستويات من A إلى H تحتوى على نفس الاختبارات الفرعية التسعة المتجمعة فى 3 بطاريات.

1- بطارية لفظية تصنيف لفظى. إكمال الجمل. متشابهات لفظية.

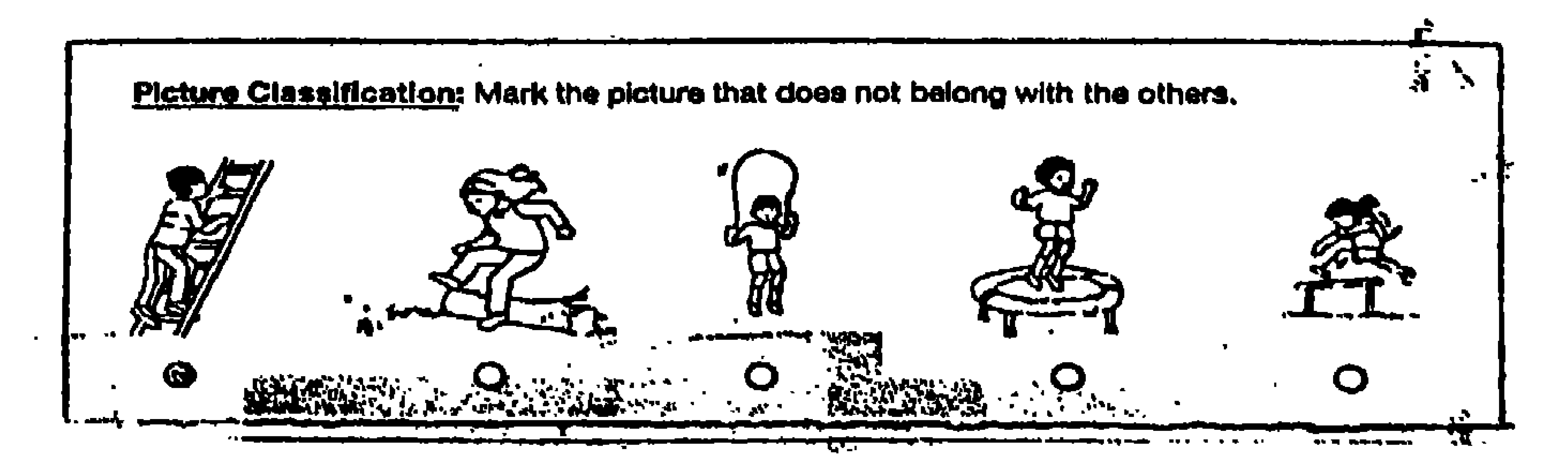
2- بطارية كمية علاقات كمية. تسلسلات رقمية. بناء معادلة.

3- بطارية غير لفظية تصنيف الأشكال. متشابهات الأشكال. تحليل الأشكال.

هذه الاختبارات الفرعية لا تستخدم كلمات أو أرقا ملكن عناصر جبرية أو شكلية فقط فكل اختبار فرعي يسبق بتمارين ممارسة مع تقديم تفسيرات مفصلة. الشكل (10- 4) يوضح موضوع نموذجى من 6 إلى 9 اختبارات فرعية بتعليمات مكدسة ومصاغة بطريقة أخرى. نجد أن الثلاث بطاريات قد أعطيت لك لطفل فى 3 دورات تعليمية كما نجد أن البطاريات الغير لفظية لا تتنبأ بالتحصيل الدراسى كما تفعل البطاريات اللفظية والكمية كما أن مقارنة أداء الأفراد على 3 بطاريات يزودنا بمعلومات مفيدة عن القدرات الخاصة أو الإعاقات.

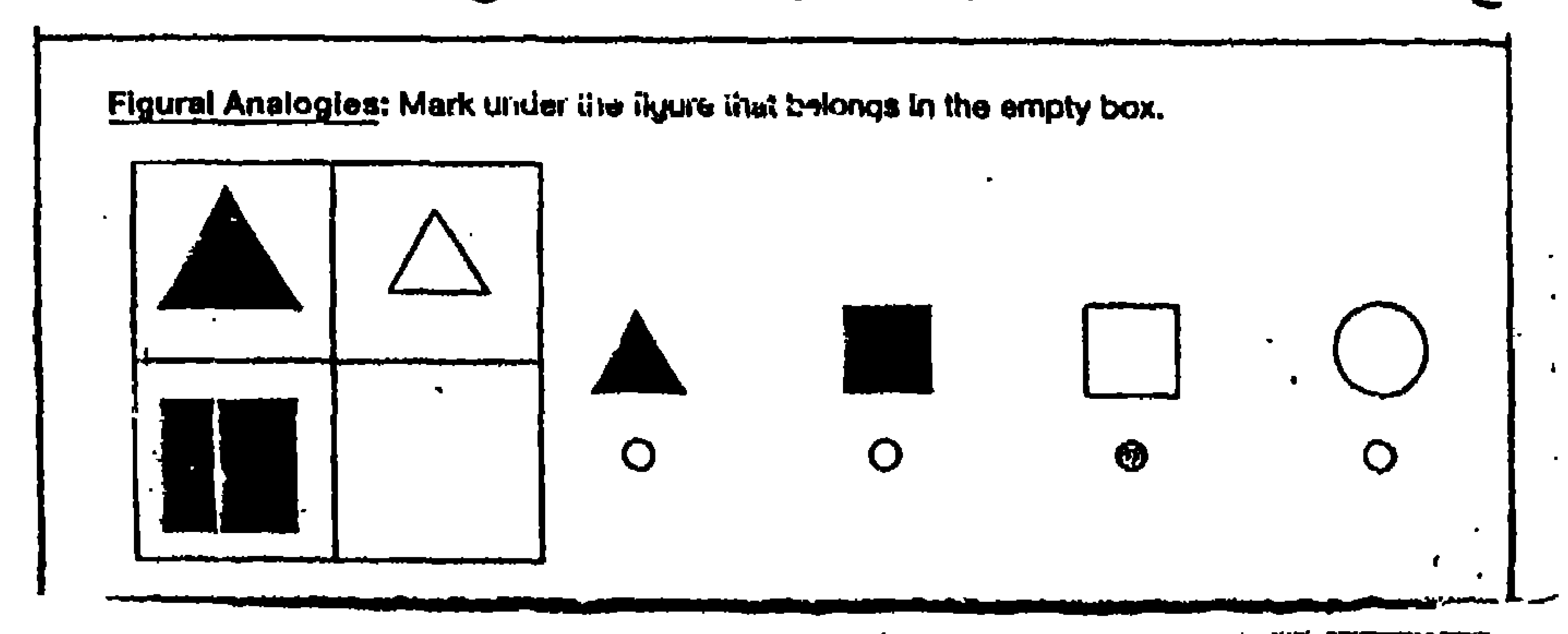
1- تصنيف الصور:

ضع علامة على الصورة التى لا تنتمى للآخرين.

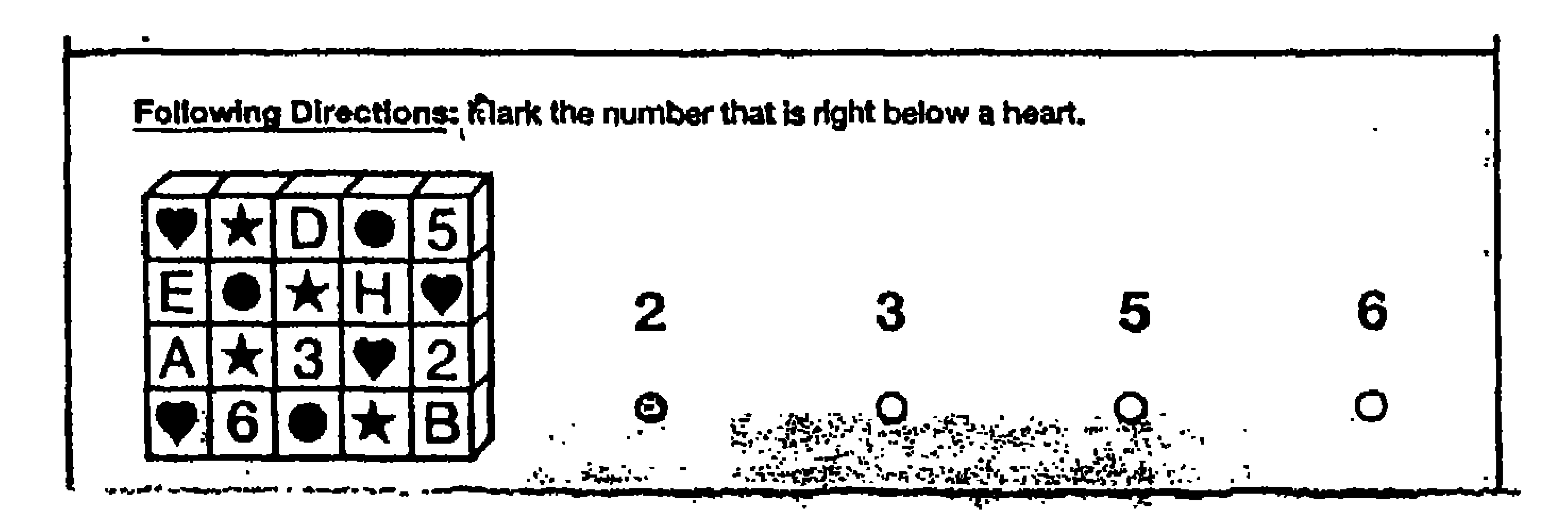


2- تشابهات الأشكال:

ضع علامة تحت الشكل الذى ينتمى إلى الصندوق الفارغ.



3- الاتجاهات المتتابعة:



ضع علامة على الرقم الصحيح أسفل القلب.

4- تسلسلات الصور:

Picture Series: Mark under the picture that shows what belongs in the empty box.

ضع علامة تحت الصورة التى توضح ما الذى ينتمى إلى الصندوق الفارغ

شكل (10- 3) بعض مفردات العينة تستخدم فى اختبار القدرات المدرسية لأوتيس ولينون (OLSAT المستوى A)

المستويات الأعلى للبطاريات متعددة المستويات تكون مناسبة لطلاب المدارس الثانوية والتى لا تختلف عن المستويات التى صممت من أجل الطلاب فى المدرسة الابتدائية ما عدا فى مستوى الصعوبة فنجد أن هذه المستويات تكون مناسبة للاختبار العام ومجموعة البالغين الغير منتقاة لأغراض متنوعة. محتوى الاختبارات فى هذا المستوى يمكن أن تشرح بواسطة مفردات من المستويات العليا لاختبار المهارات المعرفية (TCS/2).

كل مستوى من مستويات هذه البطارية يحتوى على 4 اختبارات:

1- التسلسلات: فهم وتطبيق قاعدة أو أساس فى نموذج أو تتابع الأشكال والحروف والأعداد.

2- المتشابهات: تعريف العلاقة بين زوج من الصور. إكمال زوج أخر يظهر نفس العلاقة ويمثل مشاهد للناس ، الحيوانات ، الأشياء ، الرموز البيانية.

3- الاستدلال اللفظى: تم اختباره بواسطة التنوع فى المفردات حيث تشمل تعريف العناصر الساسية للأهداف والمفاهيم وتقسيم الأهداف طبقا للعوامل المشتركة واستنتجا العلاقات من مجموعة الكلمات.

4- المذاكرة: تعريف مجموعة من الكلمات الاصطناعية (المقررات الدراسية التى لا معنى لها) تكون موجودة من اجل الدراسة ويتم استدعائها بعد حوالى 25 دقيقة اثناء القام باختبارات أخرى.

يوجد اختبار ممارسة يتم إدارته يوم أواثنين قبل جلسة الاختبار الحقيقية. يتم شرح 3 من 4 أنواع من المفردات التى تظهر فى الشكل (10- 5). فى هذه البطارية نفس أنواع المفردات من التسلسلات والمتشابهات واختبترات الاستدلال اللفظى تستخدم من المرحلة الرابعة وتستمر حتى المرحلة الـ 12 ونجد أن نفس مفردات العينة تشمل كل هذه المستويات لكن المستويين الأعلى يماثلون درجات المدرسة الثانوية ويختلفوا على أساس الصعوبة العليا المنشأة إمبريقيا للمفردات.

من الصفات المميزة لبطارية (TCS/2) انها تشمل اختبارات الذاكرة كما إنها تحتوى على تقديرات منفصلة لكلا من الاختبارات اللفظية وغير اللفظية واختبارات الذكاء بذلك نجد أن هذه المجالات الخاصة بالاستعدادات تحدد من خلال التحليل العاملى التمهيدى الذى يوجه نحو التنمية وانتقاء مفردات الاختبار. إن بطارية (TCS/2) تستخدم نظرية البند والاستجابة IRT (انظر الفصل 7) فى كلا من تنمية وتقدير الاختبار لذلك من الممكن أن نحصل على الدرجات التى لا تعكس فقط عدد المفردات التى تم اجتيازها ولكنها تعكس مستوى صعوبة المفردات الفردية التى تم اجتيازها.

1- التصنيف اللفظى: فكر كيف تكون هذه الكلمات ذات اللون الأسود متشابهة. ثم أوجد الكلمة من الخط الأسفل التى تكون متشابهة مع هذه الكلمات.

عطوف صداقة مساعد

(1) قادر على (2) نشط (3) كريم (4) جميل (5) قوى

2- التشابهات اللفظية: فى هذا الشكل كيف يكون أول كلمتين مترابطتين. ثم أوجد الكلمة التى تكمل الزوج الثانى بنفس الطريقة.

سفينة ميناء عربة نقل

(1) يسوق (2) طريق سريع (3) جراج (4) البنزين (5) قطار شحن

3- التسلسلات الرقمية: فى هذا الشكل قاعدة استخدمت لترتيب تسلسلات الأرقام. اختار الرقم الذى يجب أن يكون التالى:

3	2	1
3	2	1

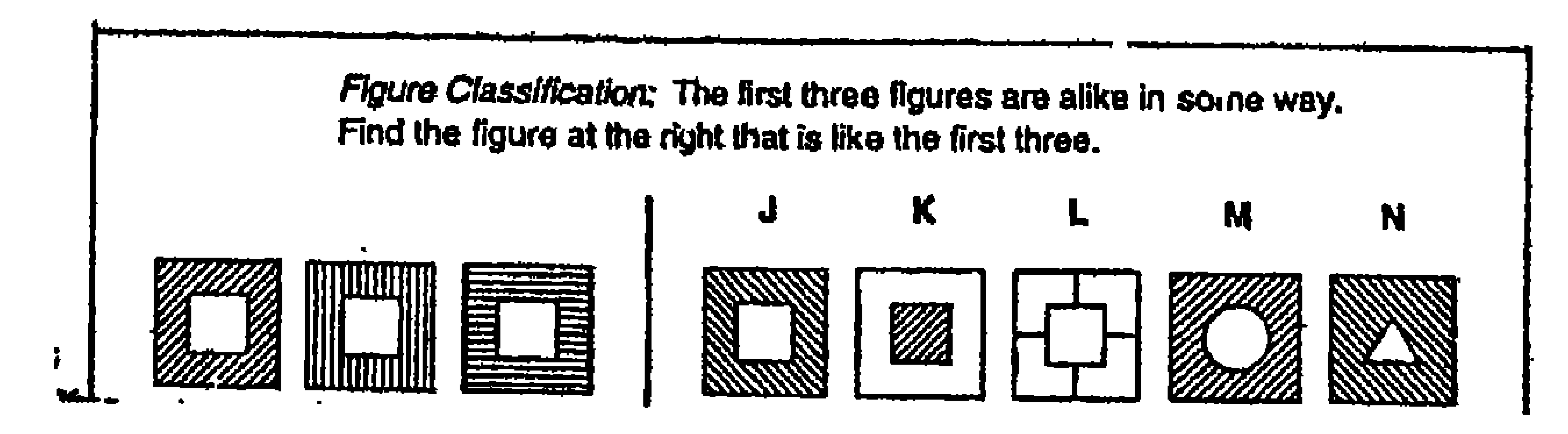
(1) 0 (2) 1 (3) 2 (4) 3 (5) 4

4- بناء المعادلات: الأرقام والعلامات التى توجد فى القمة يمكن أن تستخدم معا لكى تعطى إجابة من الإجابات بأسفل ض علامة على الإجابة الصحيحة.

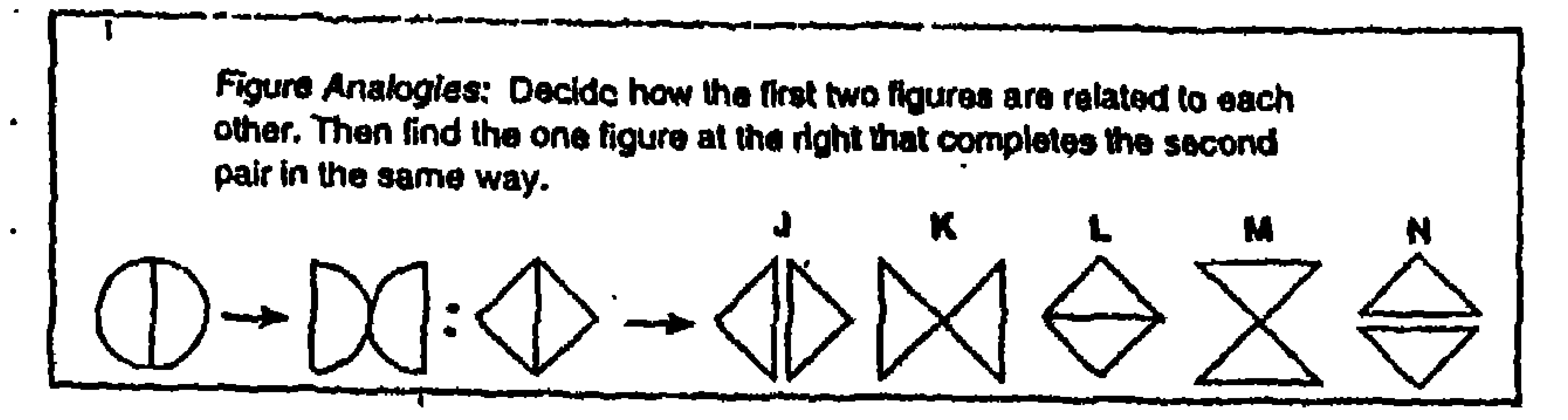
2	4	8

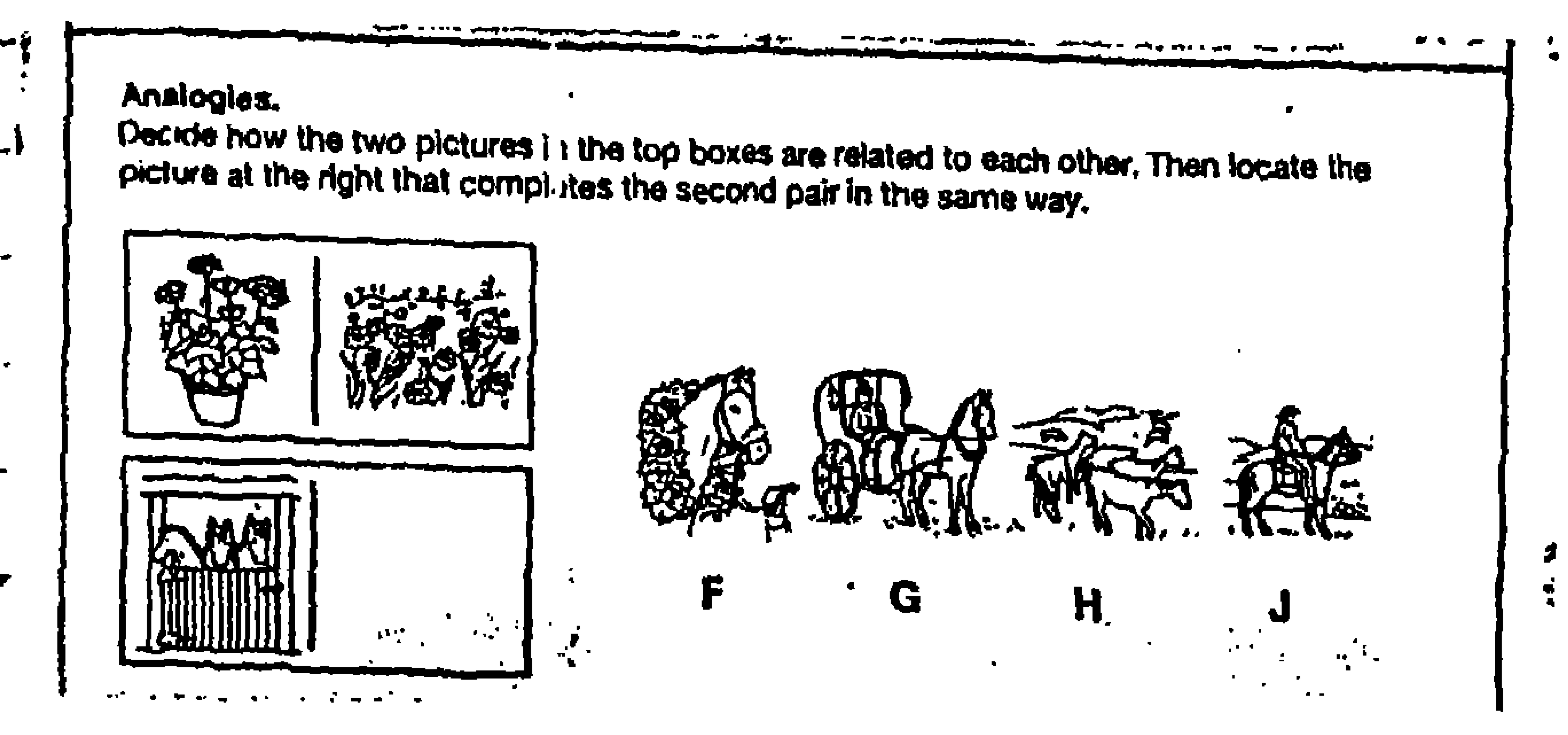
(1) 0 (2) 2 (3) 4 (4) 6 (5) 10

5- تصنيف الأشكال: أول ثلاثة أشكال متشابهة فى بعض الطرق أوجد الشكل فى اليمين الذى يشبه أول ثلاثة.



6- تشابهات الأشكال: قرر أولا كيف يكون أول شكلين لهم علاقة بكل منهم الآخر ثم أوجد الشكل على اليمين الذى يكمل الزوج الثانى بنفس الطريقة.

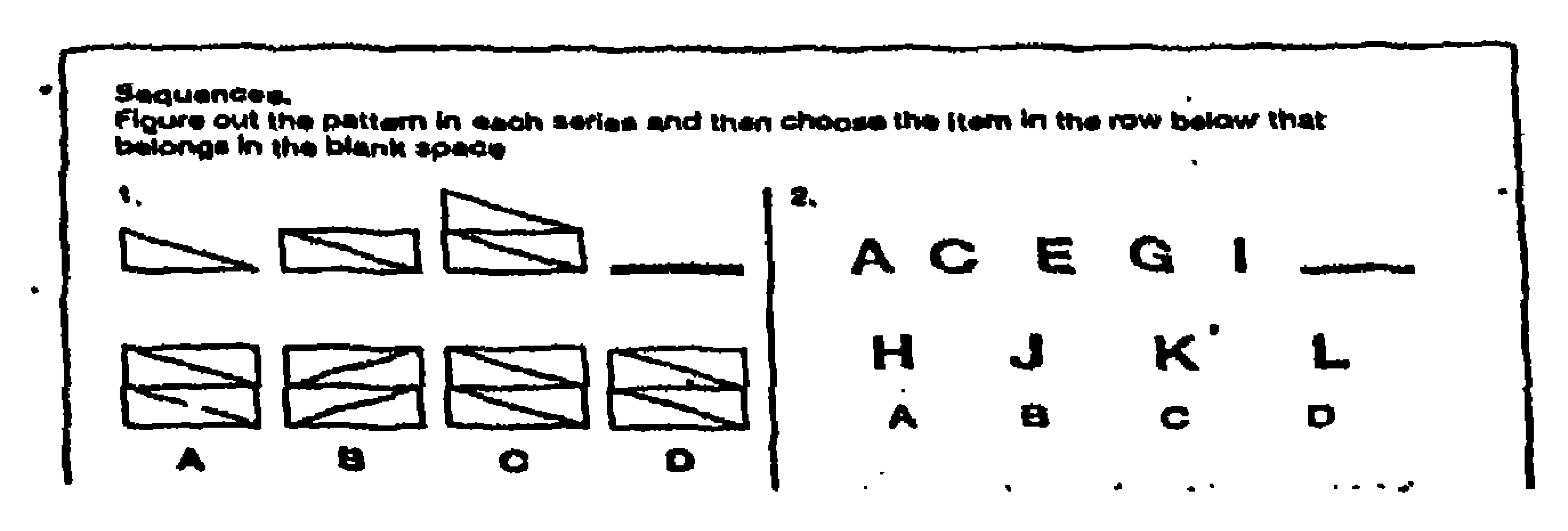




شكل (10- 4) يوضح بعض أنواع المفردات من اختبار القدرات المعرفية. الإجابات تكتب فى ورقة إجابة منفصلة

الإجابات الصحيحة هى 3، 3، 4، 2، 1، 2

1- التسلسلات: فى هذا النموذج تسلسل (شكل 10- 5) اختار الموضوع من الصف الأسفل الذى ينتمى إلى الفراغ الأسود.



2- المتشابهات: قرر كيف تكون الصور الموضحة فى الصناديق العليا مترابطة ثم حدد موقع الصورة فى اليمين التى تكمل الزوج الثانى بنفس الطريقة.

3- الاستدلال اللفظى:

1- انظر إلى الكلمة التى يوجد تحتها خط الحروف الهجائية أى من الكلمات الموجودة بأسفل تسمى شيئا يجب أن يكون دائما جزء منها الحروف الهجائية.

(1) كلمات (2) حروف (3) ارقام (4) جمل.

2- المعلومات الموجودة فى الجملتين بالأعلى ، اى من الجمل التى توجد أسفل يجب أن تكون صحيحة.

- بيج بن ساعة فى بريطانيا.
- جودى زار بيج بن.

1- الناس غالبا يزورون بيج بن.

2- عديد من الساعات فى انجلترا تكون كبيرة.

3- بيج بن اسم شخص.

4- جودى يوجد فى انجلترا.

شكل (10- 5) بعض أنواع المفردات التى استخدمت فى اختبار المهارات المعرفية

الإجابات تكتب فى ورقة إجابة منفصلة

معرفة الاستعدادات المتعددة:

كما يتضح من الرسومات البيانية المقترحة فى هذا الفصل أن هناك ميول واضحة نحو الربط بين اختبار القدرات العامة وقياس الاستعدادات المستقلة. هذه الميول توزازى الذى حدث مع الاختبارات الفردية فى الفصل 8. لكن فى الحالة الحاضرة نجد أن البطاريات متعددة المستويات الأولى صممت كنسخة جماعية لاختبارات الذكاء الفردية كما إنها تمتلك هدف محدد كتقييم الاستعدادات المدرسى أو القراءة من أجل مستويات تعليمية ناجحة. لقد أصبح من الواضح أن الدرجات الكلية الفردية يمكن أن تزود بدرجات أخرى على نطاق ضيق. هذا التطور يحدث فى كلا من الثلاث بطاريات التى نوقشت فى هذا الفصل.

فى الـ OLSAT (بطارية اختبارات القدرات المدرسية) تم الإشارة إلى أن الدرجات الكلية تكون محددة بمجموعة من الاستعدادات التعليمية اللفظية حيث نجد أن البطارية لا تصنع محاولات لتقييم الجزء الميكانيكى العملى للذكاء العام لكن هناك تقسيم لتحديد درجات الجزء اللفظى والغير لفظى. هذا التمييز يركز أولا على مفردات الاختبار التى تتطلب أولا استخدام اللغة فى إجابات الاختبار لكننا نركز أولا على استخدام الاختبار مع الطلاب الذين يمتلكون معرفة محدودة فى اللغة. من الممكن أن نحصل على درجات عن (العمر أو المرحلة) من خلال خمس اختبارات تشمل (الفهم اللفظى ، العمر أو المرحلة) من خلال خمس اختبارات تشمل (الفهم اللفظى ، الاستدلال اللفظى ، الاستدلال التصورى ، الاستدلال الشكلى ، الاستدلال الكمى) وهى تعبر عن التصنيفات اللفظية والغير لفظية من خلال المقارنة بين الأداء لنسبى للأفراد فهذا يساعد على تحديد قوة الطلاب وضعفهم فى هذه الاختبارات.

اختبار القدرات المعرفية (COGAT) يزودنا بمعايير ليس فقط من أجل الدرجة الكلية لكن أيضا من أجل درجات الاستدلال اللفظى والكمى والغير لفظى. التقارير عن الدرجات الفردية يجب أن يشمل رسوم بيانية توضح الدرجات

الثلاثة بجانب الدرجة الكلية ولكن يجب أن يتم الاهتمام بالرسم البيانى الجانبى من أجل التنبؤ بالتحصيل الدراسى (See e.g Riverside 2000, 1994, p. 44) بذلك نجد أن معرفة الاستعدادات المتعددة تحركت ضمن خطوة جديدة للأمام.

اختبار المهارات المعرفية (TCS/2) يوضح لنا قيمة البيانات الخاصة بالاستعدادات المتعددة فى تقييم الأداء الدراسى. المعايير المتاحة لا تستخدم فقط من أجل الدرجات الكلية للبطارية لكن أيضا من أجل الاختبارات الفرعية الأربعة (المتشابهات والتسلسلات) ومن أجل الاختبارات الغير لفظية. لقد اهتمت البطارية بتقييم السمات المعرفية الثلاثة التى تم تحديدها خلال بحث التحليل العاملى وهى الاستدلال اللفظى والاستدلال الغير لفظى والذاكرة.

قياس الاستعدادات المتعددة:

اختبارات الذكاء التقليدية إذا تم إدارتها فرديا أو جماعيا فقد صممت أولا لتمنح قياس عالى لمستوى عام للتنمية المعرفية مثل نسبة الذكاء IQ نجد أن كلا من التنمية النظرية والعملية جذبت الانتباه للاستعدادات المختلفة المحددة بواسطة اختبارات الذكاء هذه التنمية تؤدى إلى بناء اختبارات منفصلة لقياس الاستعدادات القابله للتطبيق على نطاق واسع وأيضا تؤدى إلى الاهتمام بماذا تقيس اختبارات الذكاء.

عدد من الأحداث تساهم فى الاهتمام بقياس الاستعدادات المختلفة أولا: هناك معرفة متزايدة للتنوع الفردى فى الأداء على اختبارات الذكاء كما ظهرت محاولات غير حاسمة لتقارن بين الأفراد من خلال الاختبارات الفرعية المختلفة التى سبقت بطاريات الاستعدادات المتعددة بالعديد من السنوات فالاختبارات الفرعية كانت غير موثوق فيها لتقارن بين الأفراد بطريقة عادلة لذلك عند بناء

اختبارات الذكاء تم اختيار الاختبارات الفرعية التى تترابط بطريقة قليلة مع بعض أجزاء المقياس يتم استثنائها كما ظهرت مجهودات لتقلل من التنوع الفردى فى الأداء. بسبب الطريقة التى أنشأت بها معظم اختبارات الذكاء فإن الإداء على هذه الاختبارات يمكن أن يكون صنف إلى أكثر من تصنيفين كالاختبار اللفظى وغير اللفظى أو اللغوى والكمى.

يتم تنمية البطاريات ذات الاستعدادات المتعددة بواسطة اختبارات الذكاء العامة وقد أصبح من الواضح أن هذه الاختبارات تكون أساسا أوليا لقياس الفهم اللفظى كما أن هناك مجالات محددة مثل القدرات الميكانيكية كانت غير ملموسة ما عدا فى بعض الأداءات والمقاييس الغير لغوية ومثل هذه الحدود لاختبارات الذكاء أصبحت واضحة لدى علماء النفس الذين قاموا بتحديد نوعية مصطلح الذكاء وقد قام بعض علماء النفس بالتمييز بين الذكاء الدراسى والعملى وهناك أخرون تحدثوا عن الذكاء المجرد والميكانيكى والاجتماعى. لقد أنشأت اختبارات الاستعدادات الخاصة لتدعم اختبارات الذكاء لكن التحليل عن قرب يوضح لنا أن اختبارات الذكاء يمكن أن تقيس بعض الاستعدادات الخاصة مثل القدرة اللفظية والرقمية.

من مميزات اختبارات الاستعدادات الفارقة إنها زودت بالنشاطات المختلفة لعلماء النفس فى الإرشاد المهنى وفى انتقاء وتصنيف الأشخاص على الوظائف العسكرية والصناعية. التنمية المبكرة للاختبارات الخاصة فى المجالات الكتابية والميكانيكية والمجالات الأخرى المهنية تكون انعكاس لهذه الاهتمامات كما يتم استخدام بطاريات الاختبار من أجل انتقاء المتقدمين لوظيفة معينة للقبول بمدارس الطب والحقوق والهندسة وطب الأسنان والمجالات الأخرى المتشابهة وهذه البطاريات تكون فى تقدم كبير عبر السنين وهناك العديد من بطاريات الاستعدادات الفارقة مثل البطاريات التى يتم إعدادها بواسطة خدمات الجيش

وبواسطة الولايات المتحدة لخدمة العماله وقد كانت النتيجة المباشرة للانتقاء المهنى أو تقسيم العمل.

فى النهاية ، إن دراسة تنظيم السمات من خلال أساليب التحليل العاملى قد زودت بالأساس النظرى لبناية بطاريات الاستعدادات المتعددة ومن خلال البحث العاملى التحليلى للقدرات المختلفة المجمعه تحت الذكاء فإنها تكون معفة ومحددة بطريقة منتظمة ويتم انتقاء الاختبارات التى تمثل افضل قياس لواحد من العوامل أو السمات المعرفة بواسطة التحليل العاملى.

اختبارات الاستعدادات الفارقة DAT:

اختبارات الاستعدادات الفارقة من أكثر بطاريات الاستعدادات المتعددة استخداما وقد نشرت أولا فى عام 1947 وقد أجرى عليها العديد من التعديلات (Qthed formc, 1992) وقد صممت هذه البطارية من أجل إرشاد الطلاب فى التعليم وفى حياتهم المهنية من المرحلة 8 إلى 12. الطبعة الخامسة للـ DAT عبارة عن مستويين:

المستوى الأول: صمم من اجل طلاب المراحل من 7 إلى 9 وأيضا البالغين الذين أنهوا سنوات الدراسة.

المستوى الثانى: صمم من اجل طلاب المراحل من 10 إلى 12 وأيضا البالغين الأكثر من 9 سنوات دراسة الذين ربما لم يتخرجوا من المدارس الثانوية.

الـ DAT تشمل ثمانية اختبارات هى الاستدلال اللفظى ، الاستدلال الرقمى ، الاستدلال المجرد ، سرعة الإدراك الحسى ، الاستدلال الميكانيكى ، العلاقات المكانية ، الهجاء ، استخدام اللغة. فى الشكل (10- 6) يوضح لنا بعض المفردات البسيطة لـ 4 من هذه لاختبارات. كما أن هناك اختبار ممارسة يشمل

المجالات الثمانية يتم استخدامه مع المفحوصين. هناك شكل خاص لاختبارات الاستعدادات الفارقة التى تساعد فى التوجيه الشخصى والمهنى وهو اختصار الاختبارات الثمانية وطبعها فى كتيب منفصل وبهذا الشكل نجد أن هناك اختبارات ممارسة يتم اختيارها من أجل المناسبين لمهن محددة.

الـ DAT مثل معظم الاختبارات السائدة التى يستخدم معها الكمبيوتر ولقد حدث تطور كبير للاختبارات مثل ظهور الاختبار التكيفى المحسوب CAT الذى يشكل (التكيفى DAT) المتاح منذ عام 1987. نجد أن المفحوص يأخذ المفردات المناسبة لمستوى أدائه كما فى كل الاختبارات التكيفية لذلك نجد أن النسخة الخاصة بـ CAT والمفردات البدائية المستخدمة فى DAT ثم تحليلها بواسطة نموذج Rasch (راش) وهو نموذج مبسط لنظرية البند والاستجابة (انظر الفصل 7).

1- الاستدلال اللفظى: اختار زوج الكلمات الصحيحة لتملأ الفراغ الأسود. أول كلمة للزوج توضع فى الفراغ الأسود فى بداية الجملة والكلمة الثانية للزوج توضع فى نهاية الجملة .... يكون لها زعنفه والطائر يكون له ....

(1) مياه . ريش. (2) سمكة . جناح.

(3) القرش . عش. (4) زعنفه الحوت . يطير.

(5) سمكة . سماء

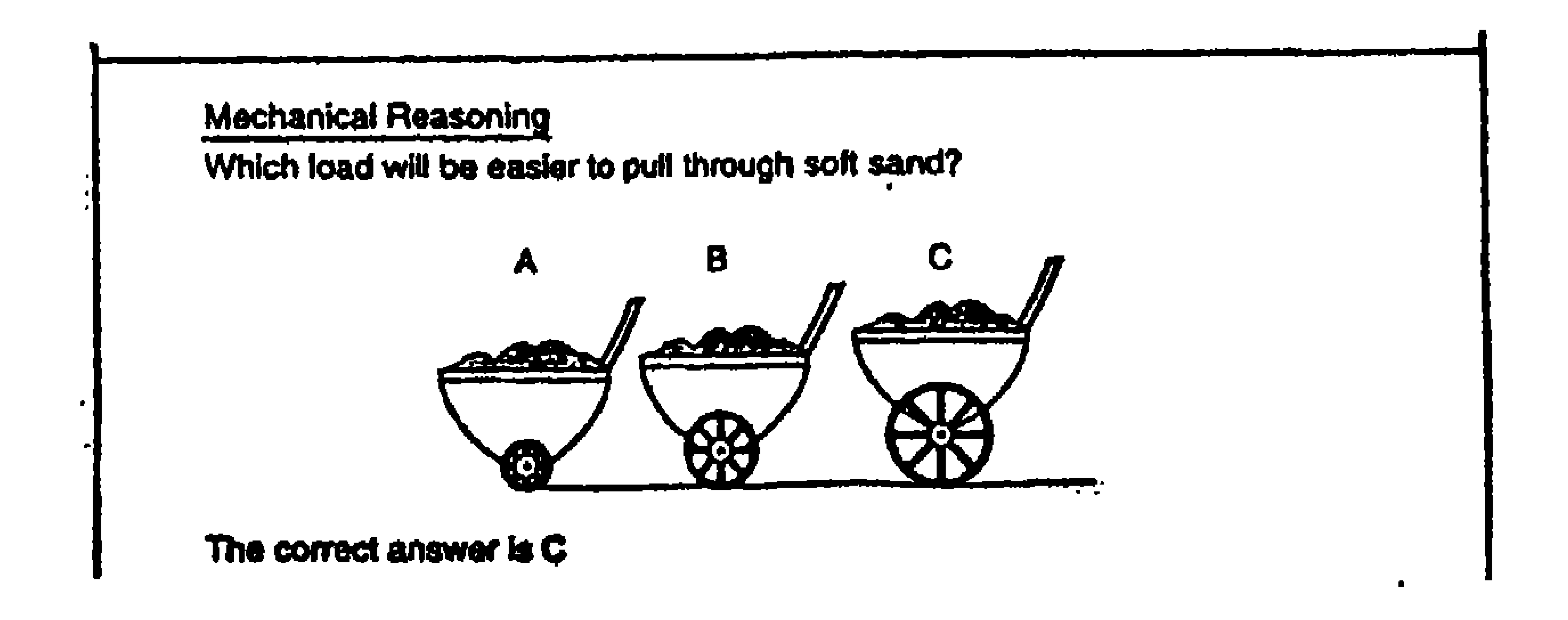
الإجابة الصحيحة تكون (2).

2- الاستدلال الرقمى: ما هو الرقم الذى يحل محل R فى هذه المشكلة من خلال الجمع.

R 7	1	2	3	4	5
R =	9	6	4	3	غير ذلك
88					

الإجابة الصحيحة تكون (1).

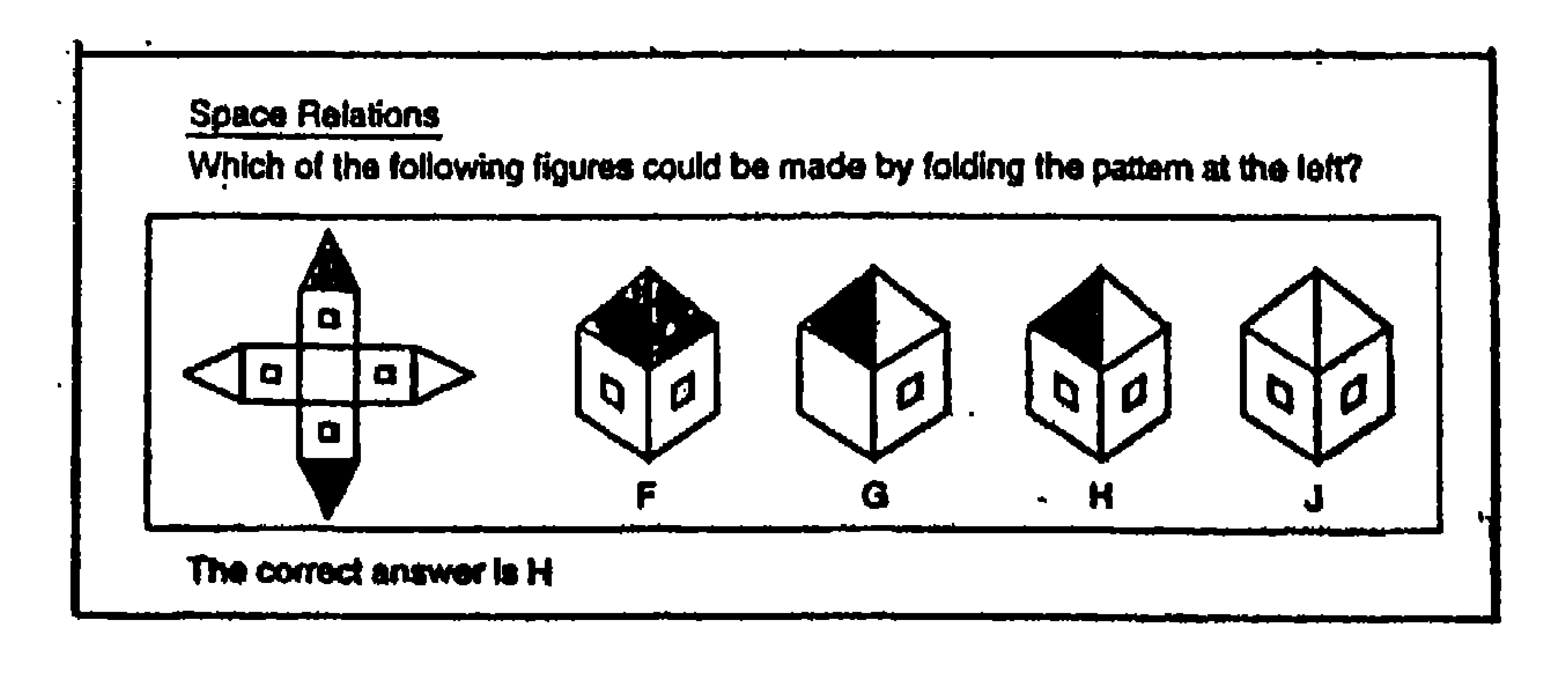
3- الاستدلال الميكانيكى: اى جمل يكون أسهل لسحبه خلال الرمل.



الإجابة الصحيحة تكون (3).

4- العلاقات المكانية: أي من الأشكال التالية يمكن أن يصنع بواسطة ثنى النموذج من الشمال.

الإجابة الصحيحة تكون (3).



شكل (10- 6) عينة من مفردات اختبارات الاستعدادات الفارقة فى الطبعة الخامسة

حتى يتم التأكد من صدق DAT ثم تجميع البيانات الخاصة بها على مدار سنتين بواسطة كل من ناشر الاختبار وبواسطة الباحثين المستقلين الذين طبقوا البطارية على مجموعات مهنية مختلفة أو فى مشاريع البحوث. معظم هذه البيانات تكون مختصة بالصدق التنبؤي لتحصيل المدارس الثانوية فى كلا من البرامج المهنية والدراسية ونجد أن الترابطات تكون مرتفعة حتى مع وجود فواصل طويلة كثلاث سنوات بين بيانات الاختبار وبيانات المحك. على الرغم من أن النتائج تكون مشجعه بطريقة قليلة لكن يجب أن نأخذ فى الاعتبار التنبؤات المختلفة حيث نجد أن الاختبارات اللفظية تترابط بطريقة عالية مع الدورات الانجليزية وكذلك الاختبارات الرقمية تترابط بطريقة عالية مع الدورات الرياضية لذلك هناك دليل على وجود عامل عام كبير للأداء فى كل عمل دراسي (على سبيل المثال الاستدلال اللفظي يرتبط ارتباط عالي مع معظم الدورات) بذلك نجد أن درجات الاستدلال الرقمي + درجات الاستدلال اللفظي تقدم كمؤشر للاستعداد المدرسى كما أن حساب درجات الاختبارات الفرعية (الاستدلال الرقمي والاستدلال اللفظي) يستخدم كمؤشر للترابط بنسبة 0.70 ، 0.80 مع محك التحصيل الدراسي. ولقد أعدت معايير مستقلة لكل واحد من هذه الدرجات ويتم رسم بروفيل للـ DAT (انظر شكل 4- 6) كما أن هناك بطارية جزئية تحتوى على الاختبارات الفرعية للاستدلال اللفظي والاستدلال الرقمي التى تستخدم كمؤشر عام للاستعداد المدرسى.

من زاوية أخرى هناك دليل على أن الاختبارات التقليدية للذكاء العام أو الاستعداد المدرسى إذا صممت من أجل الإدارة الجماعية أو الفرعية تحصد معاملات صدق جوهرية مع تنوع فى المحكات التعليمية والمهنية (L.S.

Cottfredson, 1986, Guion & Gibson 1988, Pearlman, *et al*, (1980, Schmidt, Huntor Pearlman & Shane 1979 هذه الاختبارات تشمل نفس تجميع المهارات المعرفية والمعرفة المقيمة بواسطة درجات الاستدلال الرقمي + الاستدلال اللفظي من الـ DAT. لذلك يمكن أن نرى الفجوة بين الذكاء واختبارات الاستعدادات المتعددة التي تم ربط نهايات كلا منهما. اختبارات مثل الـ DAT تم وضع تأكيد كبير على استخدامها وتفسيرها للدرجات مثل (الاستدلال الرقمي + الاستدلال اللفظي) وفي نفس الوقت الاختبارات العامة البدائية تؤكد على استخدام وتفسير درجات الاختبارات الفرعية المستقلة وتحليل البروفيلات. لإعادة النظر في DAT انظر N.Schmitt (1990); Hottrup (1990).

بطارية الاستعدادات متعددة الأبعاد MAB:

النموذج المنشأ لاختبار القدرة قد ضرب به مثلا في أكثر بطاريات الاستعدادات متعددة الأبعاد لقد نشرت هذه البطارية عام 1984 وقد صممت كاختبار جماعي يقيم نفس الاستعدادات مثل مقياس فيسشلر Wechsler لذكاء الراشدين الذي عدل بواسطة WAIS-R (انظر الفصل 8). الـ MAB تشمل خمس اختبارات فرعية في المقياس اللفظي وخمس في مقياس الأداء ومن خلالها تجمع نسبة الذكاء الانحرافة للمقياس اللفظي والعملي والمقياس الكلى. بذلك نجد ان MAB تكون مناسبة للمراهقين والراشدين لكنها لا تستخدم مع الأشخاص المتخلفين عقليا.

الاختبارات الفرعية العشرية تمتلك نفس أسماء الاختبارات الفرعية المناظرة WAIS-R ما عدا في شئ واحد وضع في الجدول بأسفل.

الأداء	لفظي
رموز الأرقام	معلومات
إكمال الصور	الفهم
مكاني	الحساب
ترتيب الصور	المتشابهات
تجميع الأشياء	المعاني

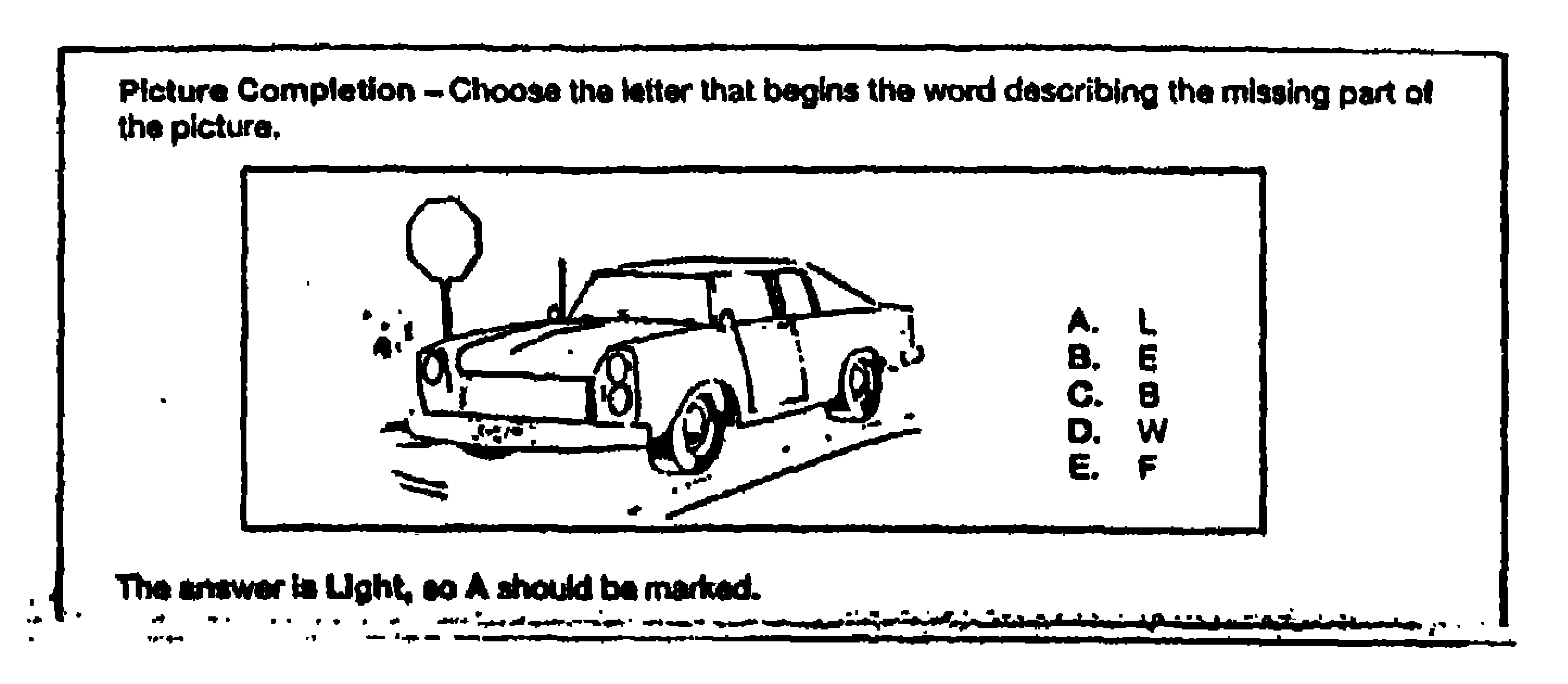
الاختبار الفرعي لرسم المكعبات لوكسلر WAIS-R استبدل بواسطة اختبار فرعى مكاني فى MAB، لقد ظهر بحث جديد يعرض تصميم مفردات الاختبار بالورقة والقلم حتى يتم قياس نفس الوظائف التى يشتمل عليها مقياس وكسلر WAIS-R الفردي. فى الشكل (10- 7) يوضح موضوع بسيط من أجل اختبار فرعى لإكمال الصورة وواد أخر من أجل اختبار فرعى مكاني حيث نجد أن مهمة المفحوصين فى الاختبارات الفرعية لإكمال الصورة هى أن المفحوص يجب أن يفكر فى اسم الجزء المفقود ثم يحدد أول حرف من هذا الاسم لكن فى الاختبارات الفرعية المكانية يختار المفحوص واحد من 5 اشكال فى اليمين يمكن أن ينتج بعد تحول بسيط الشكل الذي يوجد فى الشمال.

الخمسة اختبارات الفرعية اللفظية توجد فى كتيب والخمسة اختبارات الفرعية الغير لفظية توجد فى كتيب ثاني وكل كتيب يبدأ بمشكلات تدريبية تشرح أنواع المفردات المستخدمة فى ثلاث من خمس اختبارات فرعية وكل اختبار فرعى يبدأ بواحد او اثنين أو ثلاث مفردات موضحه بطريقة كبيرة كما أن كل اختبار فرعى له تعليمات عامة وتعليمات محددة هذه التعليمات قد توجه شفهيا أو سمعيا عن طريق جهاز تسجيل أو قد تتم بواسطة الكمبيوتر والإجابات تسجل فى جداول إجابات منفصلة أو على الكمبيوتر وقد يتم تحديد درجات الأفراد يدويا

من الأدلة أو بواسطة الكمبيوتر لكن هناك نسخ من MAB يتم إدارتها إدارة كاملة بالكمبيوتر.

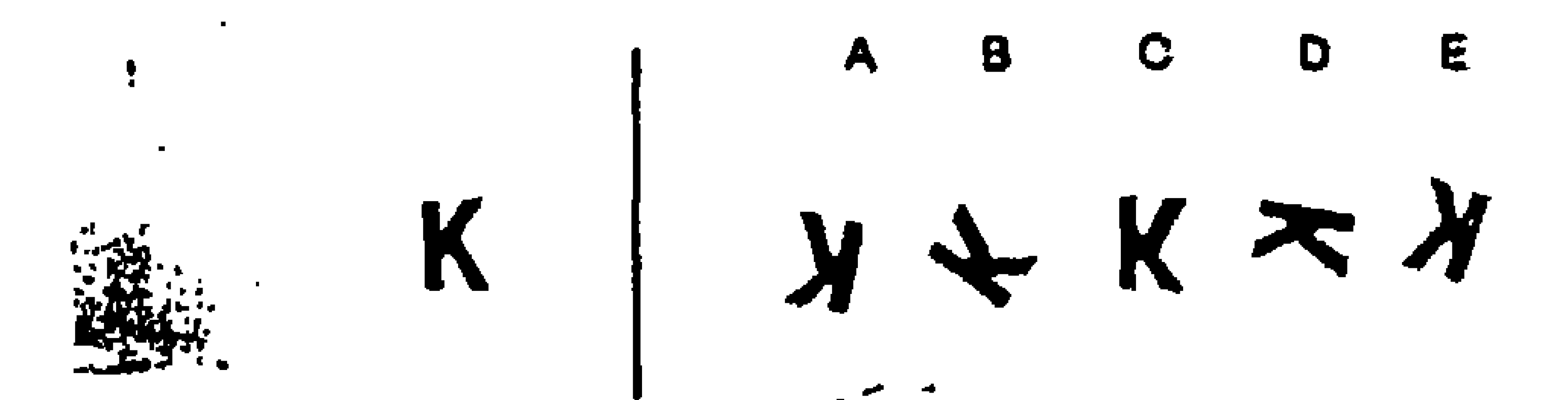
بالرجوع إلى الجدول المعيارى والدرجات المحددة لكل واحد من العشر اختبارات الفرعية يمكن أن تحول إلى متكافئات ذات الدرجة المقياسية الموحدة (الدرجات المعيارية بين المتوسط = 5 والانحراف المعيارى = 10) عند حساب درجات المقياس اللفظى والعملى والمقياس الكلى نستخدم نسبة الذكاء الانحرافية (المتوسط = 100 والانحراف المعيارى = 15) من خلال المجموعات ذات المستوى العمرى التاسع الذى يمتد من 16 إلى 74 سنة هذه الدرجات يمكن أن تستخدم فى رسم بروفيل مناسب لعمر الفرد. يفضل استخدام درجات بطارية واحدة أكثر من استخدام مجموعة مختلفة من الاختبارات لأن اختبارات البطارية كلها قد تم تقنينها على نفس العينة.

1- إكمال الصور: اختار الحرف الذى تبدأ به الكلمة التى تصف الجزء المفقود فى الصورة.



2- المكانى: اختار شكل واحد من يمين الخط العمودى الذى يكون نفس الشكل على الشمال. شكل واحد يمكن أن يتحول إلى نفس الشكل فى الشمال.

Spatial – Choose one figure to the right of the vertical line which is the same as the figure on the left. One figure can be turned to look like the figure on the left; the others would have to be flipped over.



The correct answer is A; so A should be marked. The others BCDE would have to be flipped over.

شكل (10- 7) مفردات لاختبارات الأداء فى بطارية الاستعدادات متعددة الأبعاد

الإجراءات الخاصة بالقياس النفسى التى تم إتباعها فى تنمية الـ MAB كانت ذات نوعية فنية عالية وكل خطوة فيها كانت مدعمة ببحث مكثف لتمتد على أكثر من عقد من الزمان. من الصفات الجديرة بالانتباه أن MAB تكون رابطة مبريقية لـ WAIS-R كما إنها توضح أن كل اختبار فرعى MAB ، ودرجات املقياس الكلى واللفظى والأداء تكون مترابطة بطريقة عالية مع درجات WAIS-R التى طبقت على عينة مختلفة الخواص من 145 مراهق وراشد. فى الحقيقة هذا الارتباط كان مرتفع مثل ترابط بين درجات WAIS-R ما عدا فى (الرموز الرقمية والمكانية ، اختبار رسم المكعبات للذكاء). لقد تم التوصل إلى الترابط بين MAB وWAIS-R على الرغم من أن كل مفردات MAB تكون جديدة ولم يكن هناك مفردات مشتركة بين المقياسين. لكن نجد أن الموازنة الخطية لدرجات MAB إلى درجات WAIS-R تمت من خلا عينة مختلفة الخواص لـ 160 مستجيب فى عمر 16- 35 سنة وقد تمت الموازنة على عينة من الطلاب وعينات من المرضى النفسيين وقد اقترح أ نهذا التدرج كان على عينات مختلفة من المفحوصين.

من زاوية أخرى هذه البطارية ضربت مثلا يوضح الميل الشامل نحو الدرجات الهرمية حيث إنها تزودنا بدرجات كلية للاختبارات الفرعية العشرة ، اختبارات الأداء ولاختبارات اللفظية وفى النهاية نحصل على درجات كلية شاملة من أجل البطارية الكاملة مستخدم الاختبار يتميز بالمرونة فى اختيار مستوى الدرجات المنسب للاختبار.

المراجع

1) أحمد عبد الخالق (1987) : قلق الموت، الكويت ، سلسلة عالم المعرفة .

2) أحمد عكاشة (1980) : الطب النفسي المعاصر ، القاهرة الانجلو المصرية .

3) آمال أحمد محروس (1988) : مقارنة عدة طرق لحساب ثبات الاختبارات مرجعية المحك . رسالة دكتوراه غير منشورة ، كلية التربية جامعة الزقازيق .

4) انتصار يونس (1978) : السلوك الإنساني ، القاهرة ، دار المعارف .

5) أنور محمد الشرقاوي ، سليمان الخضري الشيخ ، أمينة محمد كاظم ، نادية محمد عبد السلام (1996) : اتجاهات معاصرة في القياس والتقويم النفسي والتربوي ، القاهرة ، مكتبة الانجلو المصرية .

6) جابر عبد الحميد جابر (1983) : التقويم التربوي والقياس النفسي ، ط 1 ، القاهرة ، دار النهضة العربية .

7) حامد عبد السلام زهران (1978) : الصحة النفسية والعلاج النفسي ، القاهرة ، عالم الكتب .

8) سبع محمد لبدة (1985) : مبادئ القياس النفسي والتقييم التربوي ، ط 3 ، عمان ، الأردن ، جمعية أعمال المطابع التعاونية .

9) سعيد محمد بامشموش ، السيد محمد خيري ، يحي محمد عبده مهني (1985) : التقييم التربوي ، ط 2 ، الرياض ، دار الفيصل الثقافية .

10) سليمان محمد سليمان (1986) : اثر بعض طرق تقدير الدرجات علي الثبات الاختبارات لمرجعية المحك ذات الاختيار من متعدد . رسالة ماجستير غير منشورة ، كلية التربية ، جامعة الزقازيق .

11) صبري محمد الدمرداش (1986) : أساسيات تدريب العلوم ، ط 1 ، القاهرة ، دار المعارف .

12) صفوت فرج (1994) : القياس النفسي ، ط 1 ، القاهرة ، دار الفكر العربي .

13) صلاح الدين ابو ناهية (1994) : القياس التربوي،ط 1 غزة ، مكتبة اليازجي.

14) صلاح الدين علام (1986) : تطورات معاصرة في القياس النفسي والتربوي ، الكويت ، المطبعة التجارية .

15) صلاح الدين علام (1995) : الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية ، ط 1 ، القاهرة ، دار الفكر العربي .

16) عبد اللطيف فؤاد إبراهيم (1965) : في المناهج وتطبيقاتها ، القاهرة ، مكتبة مصر الفجالة .

17) غانم سعيد شريف ، حنان عيسي سلطان (1981) : أساسيات القياس والتقويم في التربية والتعليم ، الرياض ، دار العلوم للطباعة والنشر .

18) فاروق عبد الفتاح (1990) : القياس النفسي والتربوي للأسوياء والمعوقين ، القاهرة ، النهضة المصرية .

19) فتحي عبد الحميد (1987) : دراسة أثر بعض العوامل الخاصة بالاختبار والمفحوصين علي الأداء الناتج من الاختبارات تحصيلية ذات اختيار متعدد . رسالة دكتوراه غير منشورة ، كلية التربية ، جامعة الزقازيق .

20) فؤاد أبو حطب ، سيد عثمان (1973) : التقويم النفسي ،القاهرة ، الانجلو المصرية .

21) فؤاد أبو حطب ، أمال مختار صادق (1980) : علم النفس التربوي ، ط 2 ، القاهرة ، الانجلو المصرية .

22) كمال دسوقي (1974) : علم الأمراض النفسية . التصنيفات والأعراض المرضية ، بيروت ، دار النهضة .

23) محمد أبو زيد إبراهيم (1991) : المنهج المدرسي ، ط 1 ، القاهرة ، مركز الكتاب .

24) محمد رضا البغدادي (1988) : الأهداف والاختبارات بين النظرية والتطبيق في المناهج وطرق التدريس ، القاهرة ، دار المعارف .

25) محمد عبد الظاهر الطيب (1988) : دراسة لمستوي قلق الامتحان بين طلاب كليات جامعة طنطا ، القاهرة ، الهيئة المصرية العامة للكتاب ، مجلة علم النفس ، العدد السادس .

26) محمد محمد سكران (1978) : عوامل رسوب الطلاب في القسم العلمي ، رسالة ماجستير ، غير منشورة ، كلية التربية جامعة عين شمس .

27) نادية محمد عبد السلام (1981) : حول معايير التقويم ونظرة جديدة . الكتاب السنوي في التربية وعلم النفس ، المجلد السابع ، القاهرة ، دار الثقافة للطباعة والنشر .

28) Anastasi , A& Urbana ,s .(1997) : Validity Psychological Testing , Prentice – Hall , N . J. PP.113-139

29) Alelanie & Keogh (1982) : Test Wise Ness As A factor In Readiness Test Performance Of Young Mexican – American Cgildren . Journal Of Educational Psychology .

الفهرس

الجداول



مؤسسة طيبة للنشر والتوزيع
7 علام حسين-ميدان الظاهر-القاهرة
ت: 27867198 - 27876470 فاكس:27876471
محمول: 0191848808 /0103450041 / 0106242622
E-mail:TIBA_ONLINE@hotmail.com